

STIC-Biotech/ChemLib

193461

From: Chan, Christina
Sent: Wednesday, June 21, 2006 8:39 AM
To: Swope, Sheridan; STIC-Biotech/ChemLib
Subject: RE: 10/825,632

~~Please rush. Thanks Chris~~

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

WRFE

-----Original Message-----

From: Swope, Sheridan
Sent: Tuesday, June 20, 2006 7:03 PM
To: Chan, Christina
Subject: 10/825,632

Chris, May I have this rushed for an allowance?

Please search:

SID 2, oligo search, against the GenEmbl and EST databases only

Collect the top 150 hits from each search.

Sheridan Swope, Ph.D.
Primary Patent Examiner
AU 1656/Recombinant Enzymes
571-272-0943 (voice)
E02D19 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)
Helping applicants get good patents.

79280

BEST AVAILABLE COPY

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 193461

TO: Sheridan Swope
Location: REM-2B71&3C70
Art Unit: 1656
Thursday, June 22, 2006
Case Serial Number: 10/825632

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

BOB
barbara.obryen@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

Thu Jun 22 09:04:26 2006

OM nucleic - nucleic search, using sw model
Run on: June 22, 2006, 03:28:17 ; Search time 14067 Seconds
(without alignments)
12402.667 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aadgttaagctcgagg.....agaattactaaaaaaaaa 3120

Scoring table: OLIGO NUC
Gapop '60.0 , Gapext 60.0
Searched: 48236798 seqs, 27959665780 residues
Word size : 1
Total number of hits satisfying chosen parameters: 96473154

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

- Database :
- 1: gb_est1.*
 - 2: gb_est3.*
 - 3: gb_est4.*
 - 4: gb_est5.*
 - 5: gb_est6.*
 - 6: gb_hic.*
 - 7: gb_est2.*
 - 8: gb_est7.*
 - 9: gb_est8.*
 - 10: gb_est9.*
 - 11: gb_ges1.*
 - 12: gb_ges2.*
 - 13: gb_ges3.*
 - 14: gb_ges4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	2767	88.7	4535	6	BC040203
2	1664	53.3	2292	6	CR609512 full-length
3	966	31.0	2649	14	AY411615 Homo sapi
4	782	25.1	1292	6	AF175225 Homo sapi
5	769	24.6	1042	3	BM557438 AGENCOURT
6	755	24.2	884	8	CR985650 CR985650
7	726	23.3	910	2	BG479035 AGENCOURT
8	725	23.2	985	3	BQ068650 AGENCOURT
9	724	23.2	910	3	BQ675260 AGENCOURT
10	723	23.1	1265	6	AF176779 Homo sapi
11	722	22.7	7351	1	AL043338
12	708	22.4	753	1	AL040398
13	700	22.1	835	8	CR998849
14	690	22.1	830	9	DA571902
15	687	22.0	742	8	CN427229
16	686	22.0	914	4	EX390898
17	685	22.0	724	5	CK000768
18	682	21.9	820	8	CR994557
19	682	21.9	820	8	CR994557

20	676	21.7	864	1	AL542617
21	674	21.6	2649	14	AY411616
22	672	21.5	1002	7	BE733691
23	671	21.5	671	9	CX787143
24	663	21.2	804	3	BU687687
25	660	21.2	993	2	BG259714
26	652	20.9	910	4	BX372276
27	649	20.8	657	9	DA415588
28	647	20.7	789	2	BG709118
29	644	20.6	669	2	BG390103
30	641	20.5	881	8	CR990494
31	639	20.5	708	3	BU622228
32	637	20.4	722	3	BU631054
33	627	20.1	853	2	BI223892
34	626	20.1	815	4	CB990233
35	626	20.1	826	2	BI553405
36	623	20.0	695	2	BG715866
37	614	19.7	634	1	AI819365
38	607	19.5	957	3	BQ675006
39	598	19.2	612	8	CV024163
40	596	19.1	596	4	CB153590
41	596	19.1	587	9	DA190607
42	587	18.8	587	9	DA392243
43	586	18.8	586	9	DA441156
44	585	18.8	593	9	DA157274
45	585	18.7	587	9	DA159782
46	583	18.7	583	9	DA488030
47	582	18.7	582	9	DA390016
48	581	18.6	632	4	EX645431
49	580	18.6	580	9	DB168729
50	580	18.6	1041	2	BI084090
51	578	18.5	593	9	DB156844
52	578	18.5	629	9	DA140770
53	576	18.5	576	9	DA301389
54	576	18.5	576	9	DA690898
55	576	18.5	577	9	DA478474
56	576	18.5	627	2	BG752328
57	575	18.4	575	9	DA163570
58	573	18.4	573	9	DA769628
59	573	18.4	574	9	DA931907
60	571	18.3	573	9	DA858700
61	562	18.0	760	1	AI917735
62	561	18.0	563	8	DA860616
63	561	18.0	678	8	CN427235
64	558	17.9	673	7	AW303607
65	557	17.9	557	9	DA997805
66	556	17.8	558	9	DA756453
67	555	17.8	788	2	BG163397
68	553	17.7	553	9	DB074943
69	552	17.7	566	9	DA268082
70	548	17.6	553	8	CN427230
71	548	17.6	612	8	CN427227
72	547	17.5	1278	6	AF173382
73	546	17.5	546	9	DA052731
74	543	17.4	897	3	BU191638
75	541	17.3	541	9	DA809185
76	539	17.3	556	9	DB339434
77	539	17.3	540	9	DA437440
78	539	17.3	542	9	DA396556
79	539	17.3	569	9	DB199007
80	539	17.3	551	9	DA446578
81	538	17.2	537	3	BU784521
82	534	17.1	534	9	DA515446
83	534	17.1	578	9	DA927284
84	527	16.9	578	9	DA773829
85	526	16.9	674	7	BE888665
86	525	16.8	565	9	DA694456
87	525	16.8	717	9	DA828964
88	523	16.8	523	9	DR022989
89	522	16.7	546	9	DB146190
90	521	16.7	523	9	DA990713
91	521	16.6	542	8	CN427232
92	518	16.6	639	7	AW958665

Qy	261	AATCGCGGACTGTGAGGAGATATTGAATCACAGATCGGCTAAATTTGGAGCCTTTT	320	Db	2756	CATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATTATT	2815
Db	1676	AATCGCGGACTGTGAGGAGATATTGAATCACAGATCGGCTAAATTTGGAGCCTTTT	1735	Qy	1401	TATCCAGTAGAAGATGATGTTATGAAAGCGAGAGACTCATTTGAGTCACTGCTGATTC	1460
Qy	321	TGTTGAGCGGTATTTCTGAGAGTCAAGCTTAAAGAGCTGCTGCGGATACAGAAATATCA	380	Db	2816	TATCCAGTAGAAGATGATGTTATGAAAGCGAGAGACTCATTTGAGTCACTGCTGATTC	2875
Db	1736	TGTTGAGCGGTATTTCTGAGAGTCAAGCTTAAAGAGCTGCTGCGGATACAGAAATATCA	1795	Qy	1461	TGTGAGCGGACTTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCATGACAT	1520
Qy	381	TGGCTACATGATGGCTTAAGGACACACATGATTTCAATGTTTGTGAGAGGATCATCCAGA	440	Db	2876	TGTGAGCGGACTTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCATGACAT	2935
Db	1796	TGGCTACATGATGGCTTAAGGACACACATGATTTCAATGTTTGTGAGAGGATCATCCAGA	1855	Qy	1521	CTTTTCATGTTTTTCCCAAGTCAAGAGAGAAATTTGAGTTTATTTTTTGGCCTCTCAATG	1580
Qy	441	TGGACCTCATTCAGACAGATCTATTACCTTCGCATGCTGCTGAGAGACAGAGAAATAC	500	Db	2936	CTTTTCATGTTTTTCCCAAGTCAAGAGAGAAATTTGAGTTTATTTTTTGGCCTCTCAATG	2995
Db	1856	TGGACCTCATTCAGACAGATCTATTACCTTCGCATGCTGCTGAGAGACAGAGAAATAC	1915	Qy	1581	CAAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAA	1640
Qy	501	ACTGTTTTATCTGAAATTTCCCAAACTATCAATAGAGAGAGCTTTAATGCTCTCTG	560	Db	2996	CAAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAA	3055
Db	1916	ACTGTTTTATCTGAAATTTCCCAAACTATCAATAGAGAGAGCTTTAATGCTCTCTG	1975	Qy	1641	ACGATCCAGTGGTGGCTCGCTGCCCAAGTGATTTCAAGTGTCTATCAAGAGGAGAT	1700
Qy	561	GAAAGCTCTTTTGGATCTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGA	620	Db	3056	ACGATCCAGTGGTGGCTCGCTGCCCAAGTGATTTCAAGTGTCTATCAAGAGGAGAT	3115
Db	1976	GAAAGCTCTTTTGGATCTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGA	2035	Qy	1701	AGCAATTTACAGTGGTGAATGGAAGTCTTTGGCCGCGCATGATCTAATATCCAAAGTTGA	1760
Qy	621	AGAACTATTAGAGAAAGAAACGATTTGGAAACAGTCGGAATTTGCTTTCTTACGATTATCA	680	Db	3116	AGCAATTTACAGTGGTGAATGGAAGTCTTTGGCCGCGCATGATCTAATATCCAAAGTTGA	3175
Db	2036	AGAACTATTAGAGAAAGAAACGATTTGGAAACAGTCGGAATTTGCTTTCTTACGATTATCA	2095	Qy	1761	TGAAGTCAGAGGCTGGTATATTTTGAAGCGCACAAAGACTCCCTTTTGAAGCATCACCT	1820
Qy	681	CCAAGGAAGTGGAACATTTCTGTTTCAAGCGGTGTGGAATTTATCAGTAAAGATGG	740	Db	3176	TGAAGTCAGAGGCTGGTATATTTTGAAGCGCACAAAGACTCCCTTTTGAAGCATCACCT	3235
Db	2096	CCAAGGAAGTGGAACATTTCTGTTTCAAGCGGTGTGGAATTTATCAGTAAAGATGG	2155	Qy	1821	GTACGTAGTCAGATTACGTAAATCTCGAGAGGTGACAGAGCTGACCTGAGCCGTACTC	1880
Qy	741	AGGGCCACAAGGATTTACGCAACAACTTTAAGGCCCACTAGTGGAACTAGTTGTTC	800	Db	3236	GTACGTAGTCAGATTACGTAAATCTCGAGAGGTGACAGAGCTGACCTGAGCCGTACTC	3295
Db	2156	AGGGCCACAAGGATTTACGCAACAACTTTAAGGCCCACTAGTGGAACTAGTTGTTC	2215	Qy	1881	ACATTTCTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAAGTATAGTAAACAGAA	1940
Qy	801	CAACATACCGATGGATCCAAATTTATGCCCTGTGATCCAGATCGGATTTGCTTTATACA	860	Db	3296	ACATTTCTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAAGTATAGTAAACAGAA	3355
Db	2216	CAACATACCGATGGATCCAAATTTATGCCCTGTGATCCAGATCGGATTTGCTTTATACA	2275	Qy	1941	GAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAA	2000
Qy	861	TAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAAAGGAGCTCACTTATGT	920	Db	3356	GAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAA	3415
Db	2276	TGGCAACGATATTTGGATATCTAACATCGTAAACAGAGAAAGGAGCTCACTTATGT	2335	Qy	2001	AACAAAGGAAATTTTGGGCCAACATTTTGGATTCAGCAGGTCTCTTCTGACTATCTCC	2060
Qy	921	GCACATGAGCTAGCCAAACATCGAAAGAGATGCCAGATCAGCTGGAGTGCCTTTGT	980	Db	3416	AACAAAGGAAATTTTGGGCCAACATTTTGGATTCAGCAGGTCTCTTCTGACTATCTCC	3475
Db	2336	GCACATGAGCTAGCCAAACATCGAAAGAGATGCCAGATCAGCTGGAGTGCCTTTGT	2395	Qy	2061	TCAGAAATTTTCTCTTTTGAAGTACTTCTGGAATTTACATTTGATGGGATGCTCTACAA	2120
Qy	981	TCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAC	1040	Db	3476	TCAGAAATTTTCTCTTTTGAAGTACTTCTGGAATTTACATTTGATGGGATGCTCTACAA	3535
Db	2396	TCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAC	2455	Qy	2121	GCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCTGCTTATATATGCTGCTCC	2180
Qy	1041	TCCAGTGGTGGTAAATTTTGAATTTCTATACAGAAATATGATGATCTCGAGTGGGA	1100	Db	3536	GCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCTGCTTATATATGCTGCTCC	3595
Db	2456	TCCAGTGGTGGTAAATTTTGAATTTCTATACAGAAATATGATGATCTCGAGTGGGA	2515	Qy	2181	TCAGGTGCAATTTGCTGGAATTAATCGGTTTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCT	2240
Qy	1101	AATTTATCTGATACATCCCTATGTTGGAACAAAGAGGGGAGATCAATTCGTTATCC	1160	Db	3596	TCAGGTGCAATTTGCTGGAATTAATCGGTTTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCT	3655
Db	2516	AATTTATCTGATACATCCCTATGTTGGAACAAAGAGGGGAGATCAATTCGTTATCC	2575	Qy	2241	AGCCTCTAGGTTATGCTGTTGATGATAGACAAACAGGGGATCTCTGTCAACGAGGGCT	2300
Qy	1161	TAAACAGGTACAGCAAACTCTTAAAGTCACTTTTAAAGTCTCAGAAATATGATGATGC	1220	Db	3656	AGCCTCTAGGTTATGCTGTTGATGATAGACAAACAGGGGATCTCTGTCAACGAGGGCT	3715
Db	2576	TAAACAGGTACAGCAAACTCTTAAAGTCACTTTTAAAGTCTCAGAAATATGATGATGC	2635	Qy	2301	TAAATTTGAAGGGCCCTTTTAAATATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGA	2360
Qy	1221	TGAAGAGGATCATAGATGTCATAGATAGGAATTAATCAACCTTTTGAAGTCTTATT	1280	Db	3716	TAAATTTGAAGGGCCCTTTTAAATATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGA	3775
Db	2636	TGAAGAGGATCATAGATGTCATAGATAGGAATTAATCAACCTTTTGAAGTCTTATT	2695	Qy	2361	AGGATCCCAATATCTAGCTTCTCGATATGATTTTCACTTACATTTAGATCGTGTGGGCATCCA	2420
Qy	1281	TCAGAGGTTGAATATATTTGCCAGAGCTGGATGGAATCTCTGAGGGGAAATATGCTTGTTC	1340	Db	3776	AGGATCCCAATATCTAGCTTCTCGATATGATTTTCACTTACATTTAGATCGTGTGGGCATCCA	3835
Db	2696	TGAAGAGGTTGAATATATTTGCCAGAGCTGGATGGAATCTCTGAGGGGAAATATGCTTGTTC	2755	Qy	2421	CGGCTGTGCTTATGAGGAGATACCTCTCCCTGATGCAATTAATGAGAGGATCAGATATCTT	2480
Qy	1341	CATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATT	1400				

```
Db 3836 CCGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGATATCTT 3895
QY 2481 CAGGTTTCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAC 2540
Db 3896 CAGGTTTCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAC 3955
QY 2541 GGAAGCTTATATGGGTCACTCTGACAGAAATGAACAGGGCTATTACTTTAGATCTGTGGC 2600
Db 3956 GGAAGCTTATATGGGTCACTCTGACAGAAATGAACAGGGCTATTACTTTAGATCTGTGGC 4015
QY 2601 CATGCAAGCAGAAAAGTTCCCTCTGACCAAAATGGTTTACTGCTCTTACATGTTTCCT 2660
Db 4016 CATGCAAGCAGAAAAGTTCCCTCTGACCAAAATGGTTTACTGCTCTTACATGTTTCCT 4075
QY 2661 GGATGAGAATGTCCATTTTGCACATACAGTATATTTACTGAGTTTCTTTAGTGGGCTGG 2720
Db 4076 GGATGAGAATGTCCATTTTGCACATACAGTATATTTACTGAGTTTCTTTAGTGGGCTGG 4135
QY 2721 AAAGCCATATGATTTACAGATCTATCTCAGAGAGACACAGCAATAAGATTTCTGAATC 2780
Db 4136 AAAGCCATATGATTTACAGATCTATCTCAGAGAGACACAGCAATAAGATTTCTGAATC 4195
QY 2781 GGGAGACATATGAATCTGATCTTTTGGACCTAATCAAGAAACCTTGGATCAGATAT 2840
Db 4196 GGGAGACATATGAATCTGATCTTTTGGACCTAATCAAGAAACCTTGGATCAGATAT 4255
QY 2841 TGTGCTCTAAAGATGATATAATTTTGGACCTGTGTAGAACTCTCTGTATACACTGGCTA 2900
Db 4256 TGTGCTCTAAAGATGATATAATTTTGGACCTGTGTAGAACTCTCTGTATACACTGGCTA 4315
QY 2901 TTTAAACCAATGAGAGGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGTAT 4375
Db 4316 TTTAAACCAATGAGAGGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGTAT 4435
QY 2961 ACTGCGCATGTAACATCTACTCTCTGAAATAAATTTGGTGCCATGACAGGGTCTACGGTT 3020
Db 4376 ACTGCGCATGTAACATCTACTCTCTGAAATAAATTTGGTGCCATGACAGGGTCTACGGTT 4435
QY 3021 TGTGGTAGTAATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATTCCTG 3080
Db 4436 TGTGGTAGTAATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATTCCTG 4495
QY 3081 AGAGACCCAGCAATACCATAGAAATTAATAAAAAA 3120
Db 4496 AGAGACCCAGCAATACCATAGAAATTAATAAAAAA 4535
```

```
RESULT 2
CR609512
LOCUS
DEFINITION
2292 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DL005YD02 of B cells (Ramos cell line)
Cot 25-normalized of Homo sapiens (human).
ACCESSION
CR609512
VERSION
HTC; CNSLT cDNA.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 2292)
Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue
2 (bases 1 to 2292)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
```

COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pcwSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen.

FEATURES

source
Location/Qualifiers
1..2292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL005YD02"
/issue_type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pcwSPORT_6"

ORIGIN

```
Query Match 53.3%; Score 1664; DB 6; Length 2292;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 472 GCCATGCTGTGGTGAGAACAGAGAAAATACACATCTGTTTATTTCTGAAATTCCTCAAACTATC 531
Db 11 GCCATGCTGTGGTGAGAACAGAGAAAATACACATCTGTTTATTTCTGAAATTCCTCAAACTATC 70
QY 532 AATAGAGCAGCAGTCTTAATGCTCTCTTGAAGCCCTCTTTTGGATCTTTTTCAGGCAACA 591
Db 71 AATAGAGCAGCAGTCTTAATGCTCTCTTGAAGCCCTCTTTTGGATCTTTTTCAGGCAACA 130
QY 592 CTGGACTATGGAATGTATTTCTCGAGAGAAAGAACTATTAAAGAGAAAGAAACGCAATGGA 651
Db 131 CTGGACTATGGAATGTATTTCTCGAGAGAAAGAACTATTAAAGAGAAAGAAACGCAATGGA 190
QY 652 ACAGTCGGAATTTGCTTTTACGATTTACCAAGGAAGTGGAAACATTTCTGTTCACGCC 711
Db 191 ACAGTCGGAATTTGCTTTTACGATTTACCAAGGAAGTGGAAACATTTCTGTTCACGCC 250
QY 712 GGTAGTGGAAATTTATACAGTAAAGATGGAGGCCACCAAGGATTTACCCACACACCTTTA 771
Db 251 GGTAGTGGAAATTTATACAGTAAAGATGGAGGCCACCAAGGATTTACCCACACACCTTTA 310
QY 772 AGGCCCAATCTAGTGGAAACTAGTTCCTCCCAACATACCGATGGATCCAAATTTATGCCCT 831
Db 311 AGGCCCAATCTAGTGGAAACTAGTTCCTCCCAACATACCGATGGATCCAAATTTATGCCCT 370
QY 832 GCTGATCCAGATCGGATTTGCTTTTATACATACGACGATTTTGGATATCTAAATCGTA 891
Db 371 GCTGATCCAGATCGGATTTGCTTTTATACATACGACGATTTTGGATATCTAAATCGTA 430
QY 892 ACCAGAGAAAGAGAGAGACTCACTTATGTGCACATGAGCTAGCCAAACATGGAGAAAGAT 951
Db 431 ACCAGAGAAAGAGAGAGACTCACTTATGTGCACATGAGCTAGCCAAACATGGAGAAAGAT 490
QY 952 GCCAGATCAGCTGGAGTCGCTACCTTCTTCCAGAGAAATTTGATAGATATTCCTGGC 1011
Db 491 GCCAGATCAGCTGGAGTCGCTACCTTCTTCCAGAGAAATTTGATAGATATTCCTGGC 550
QY 1012 TATTGGTGGTGCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTA 1071
Db 551 TATTGGTGGTGCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTA 610
QY 1072 TATGAAGAAATGATGAATCTGAGGTGAAATTTATCATGTTACATCCCTATGTTGGAA 1131
Db 611 TATGAAGAAATGATGAATCTGAGGTGAAATTTATCATGTTACATCCCTATGTTGGAA 670
QY 1132 ACNAGGGGGCAGATTCAATCCGTTATCTTAAACAGGTACAGCAAACTCTTAAAGTCACT 1191
Db 671 ACNAGGGGGCAGATTCAATCCGTTATCTTAAACAGGTACAGCAAACTCTTAAAGTCACT 730
QY 1192 TTTTAAAGTGTGCAAAATTAATGATTTGATGCTGAAAGGAGGATCATAGATGTCATAGTAAG 1251
Db 731 TTTTAAAGTGTGCAAAATTAATGATTTGATGCTGAAAGGAGGATCATAGATGTCATAGTAAG 790
QY 1252 GAACATAATCAACTTTTGGATTTCTATTGTAAGAGGAGTTGAATATATTTCAGAGACTGGA 1311
```

Db	791	GAACATAATCAACCTTTTGGAGATTTCTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGA	850	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
Qy	1312	TGGACTCCTGAGGGAATAATATGCTTGGTCCATCTACTAGATGCTCCAGACTCGCCTA	1371	1 (bases 1 to 2649)
Db	851	TGAGCTCCTGAGGGAATAATATGCTTGGTCCATCTACTAGATGCTCCAGACTCGCCTA	910	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Qy	1372	CAGATAGTGTTCATCTCACTGGAATATTTATCCAGTAGAAGATCATGTTATGAAAGG	1431	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
Db	911	CAGATAGTGTTCATCTCACTGGAATATTTATCCAGTAGAAGATCATGTTATGAAAGG	970	Science 302 (5652), 1960-1963 (2003)
Qy	1432	CAGAGACTATTCAGTCACTGCTGATTTCTGTGAGCCACTAATATCTATGAAGAAACA	1491	14671302
Db	971	CAGAGACTATTCAGTCACTGCTGATTTCTGTGAGCCACTAATATCTATGAAGAAACA	1030	2 (bases 1 to 2649)
Qy	1492	ACAGACATCTGGATAAATATCCATGACATCTTTTTCATGTTTTTCCCAAAGTCAAGAG	1551	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Db	1031	ACAGACATCTGGATAAATATCCATGACATCTTTTTCATGTTTTTCCCAAAGTCAAGAG	1090	Direct Submission
Qy	1552	GAATTTGAGTTTATTTTTCCTCTGATGCAAAACAGGTTTCCGTCAATTTATCAAAAT	1611	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
Db	1091	GAATTTGAGTTTATTTTTCCTCTGATGCAAAACAGGTTTCCGTCAATTTATCAAAAT	1150	This sequence was made by sequencing genomic exons and ordering them based on alignment.
Qy	1612	ACATCTATTTAAAGGAAGCAAAATATTAACCATCCAGTGGTGGCTGCTGCCAAGT	1671	FEATURES
Db	1151	ACATCTATTTAAAGGAAGCAAAATATTAACCATCCAGTGGTGGCTGCTGCCAAGT	1210	source
Qy	1672	GATTTCAAGTGTCTCTATCAAAGAGGAGATAGCAATTAACAGTGGTGAATGGAGTCTT	1731	1..2649
Db	1211	GATTTCAAGTGTCTCTATCAAAGAGGAGATAGCAATTAACAGTGGTGAATGGAGTCTT	1270	/organism="Homo sapiens"
Qy	1732	GCCCGGCAATGATCTAATATCCAGTTGATGAAGTCAAGGCTCGTATATTTTGAAGGC	1791	/mol_type="genomic DNA"
Db	1271	GCCCGGCAATGATCTAATATCCAGTTGATGAAGTCAAGGCTCGTATATTTTGAAGGC	1330	/db_xref="taxon:9606"
Qy	1792	ACCAAGACTCCCTTTTAGAGTCACTCTGATAGTCACTGATGAGTAAATCTCTGGAG	1851	<1..>2649
Db	1331	ACCAAGACTCCCTTTTAGAGTCACTCTGATAGTCACTGATGAGTAAATCTCTGGAG	1390	/genes="DPP8"
Qy	1852	GTGACAAGCTGACTGACCGTGGCTACTCACTTCTGCTGATCAGTCAGTCAGCTGAC	1911	/locus_tag="HCM4283"
Db	1391	GTGACAAGCTGACTGACCGTGGCTACTCACTTCTGCTGATCAGTCAGTCAGCTGAC	1450	Query Match
Qy	1912	TTCTTTATAGTATAGTATGAAACAGAGAAATCCACACTGTGTGCCCTTTTACAGCTA	1971	Best Local Similarity 31.0%; Score 966; DB 14; Length 2649;
Db	1451	TTCTTTATAGTATAGTATGAAACAGAGAAATCCACACTGTGTGCCCTTTTACAGCTA	1510	Mismatches 1; Indels 0; Gaps 0;
Qy	1972	TCAAGTCTTGAAGATGACCAACTTGCACAAACAAAGGAATTTTGGGCCACCATTTTGGAT	2031	Matches 1016; Conservative 0;
Db	1511	TCAAGTCTTGAAGATGACCAACTTGCACAAACAAAGGAATTTTGGGCCACCATTTTGGAT	1570	QY 214 ATGCACAGCAATGGAACAGAACAGACAGCTGGGTGTGAGATATTTGAACCTGGGACTGT 273
Qy	2032	TCAGCAGGTCTCTTCTGACTATATCTCTCAGAAATTTTCTTTTGAAGTACTACT 2091	1590	DB 1 ATGCACAGCAATGGAACAGAACAGACAGCTGGGTGTGAGATATTTGAACCTGGGACTGT 60
Db	1571	TCAGCAGGTCTCTTCTGACTATATCTCTCAGAAATTTTCTTTTGAAGTACTACT 1630	1610	QY 274 GAGGAGAAATTTGAATCAGAGGATCGGCTTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 120
Qy	2092	GGATTTACATTTGATGGATGCTCTTACAGGCTCATGATCTACAGCTGGAAGAAATAT 2151	1630	DB 61 GAGGAGAAATTTGAATCAGAGGATCGGCTTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 120
Db	1631	GGATTTACATTTGATGGATGCTCTTACAGGCTCATGATCTACAGCTGGAAGAAATAT 1690	1650	QY 334 TCCTGGATCAGCTTAAAGAGCTGTTCGCGATACAGAGAAATATCATGGCTACATGATG 393
Qy	2152	CCTACTGCTGTTTCATATATGTTGTTCTCTCAGGT 2186	1670	DB 121 TCCTGGATCAGCTTAAAGAGCTGTTCGCGATACAGAGAAATATCATGGCTACATGATG 180
Db	1691	CCTACTGCTGTTTCATATATGTTGTTCTCTCAGGT 1725	1690	QY 394 GCTAAGGCCACACATGATTTTCATGTTGTAAGAGAAATGATCCAGATCGACCTCATTTCA 453
LOCUS	AY411615	2649 bp	linear	181 GCTAAGGCCACACATGATTTTCATGTTGTAAGAGAAATGATCCAGATCGACCTCATTTCA 240
DEFINITION	AY411615	Homo sapiens DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,		454 GACAGAAATCTATACCTTGGCATGCTGCTGAGAACAGAGAAATACACTGTTTATTTCT 513
ACCESSION	AY411615	genomic survey sequence.		514 GAAATTCACAAACTCATCAATAGACAGCAGTCTTAAATGCTCTCTTGGAGCCCTCTTTTG 573
VERSION	AY411615.1	GI:39767583		301 GAAATTCACAAACTCATCAATAGACAGCAGTCTTAAATGCTCTCTTGGAGCCCTCTTTTG 360
KEYWORDS	GSS.			574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAGAGAACTATTAAAGA 633
SOURCE	Homo sapiens (human)			361 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAGAGAACTATTAAAGA 420
ORGANISM	Homo sapiens			634 GAAAGAAACCGATTTGGAAACAGTCGGAATTCCTTACGATTTATCATCCAGGAAGTGA 693
				421 GAAAGAAACCGATTTGGAAACAGTCGGAATTCCTTACGATTTATCATCCAGGAAGTGA 480
				694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCATCGTAAAGAGTGGAGGCCACAGGA 753
				481 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCATCGTAAAGAGTGGAGGCCACAGGA 540
				754 TTTACGCAACAACTTTTAAAGGCCCAATCTAGTGGGAACTAGTTGTGCCCAACATACGGATG 813

```
Db 541 TTTAGCGCAACACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCTCCACATAGGATG 600
Qy 814 GATCCAAATATATCCCGCTGATCAGAGCTGGATTGCTTTTATACATAGCAAGATATT 873
Db 601 GATCCAAATATATCCCGCTGATCAGAGCTGGATTGCTTTTATACATAGCAAGATATT 660
Qy 874 TGGATATCTAACATCGTAACAGAGAAAGAGGAGACTCAGTATGTCGCAATAGCTTA 933
Db 661 TGGATATCTAACATCGTAACAGAGAAAGAGGAGACTCAGTATGTCGCAATAGCTTA 720
Qy 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGCTCCAAAGAA 993
Db 721 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGCTCCAAAGAA 780
Qy 994 TTTGATAGATATTTCTGGCTATTGGTGTGTCCTCAAGCTGAAACATCCCAAGTGTGT 1053
Db 781 TTTGATAGATATTTCTGGCTATTGGTGTGTCCTCAAGCTGAAACATCCCAAGTGTGT 840
Qy 1054 AAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTT 1113
Db 841 AAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTT 900
Qy 1114 ACATCCCTATGTTGGAACAGAGGAGGCGAGATTCCTGTTATCCTTAAACAGGTACA 1173
Db 901 ACATCCCTATGTTGGAACAGAGGAGGCGAGATTCCTGTTATCCTTAAACAGGTACA 960
Qy 1174 GCAAAATCTTAAAGTCACTTTTAAAGATGTCGAAATATGATGCTCGAAGGAAG 1230
Db 961 GCAAAATCTTAAAGTCACTTTTAAAGATGTCGAAATATGATGCTCGAAGGAAG 1017

RESULT 4
AF175225
LOCUS Homo sapiens tissue-type aorta MSTP135 mRNA linear HTC 01-AUG-2003
DEFINITION AF175225.1 GI:33338055
ACCESSION HTC
VERSION Homo sapiens (human)
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS Zhao, B., Xu, H.S., Tong, Y.K., Sheng, H., Qin, B.M., Liu, Y.Q., Liu, B.,
Wang, Y., Zhang, Q., Song, L., Gao, Y., Zhang, C.L., Ye, J., Ji, X.J.,
Liu, B.H., Lu, H., Chen, J.Z., Cai, M.Q., Zheng, W.Y., Teng, C.Y.,
Liu, Q., Yu, L.F., Lin, J., Gong, Q., Zhang, A.M., Gao, R.L. and Hui, R.T.
Disease Submission
Submitted (04-AUG-1999) Molecular Medicine Center for
Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,
Bei Li Shi Lu, Beijing 100037, P.R. China
FEATURES
Location/Qualifiers
1..1292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
142..1023
/codon_start=1
/product="MSTP135"
/protein_id="AAQ13650.1"
/db_xref="GI:33338056"
/translation="MTQLAKRNFGPPFWTQQVLFTLLKLFSLKVLKLDLHCWGCS
TSPGSGTGLERNILLXLFYGGPKVOLVNNRFGVKVFFRLNTPLFGLVYVXGIDNRGS
CPRGKLFEGAFKFMGQIEIDDOVEGLQYLASRYDFLDLRVGLHGNISGYSLSMAL
MQRSDIFVALAGAPVTLWIFPYDTGYTERMGHPDQNGYVLGSVAMQAEKFFSEPN
RLLHLHGFLENVHFHTSILLSFLVRAGKPYDLQIYQERHSIRVPESGHEYLHLL
HYLQENLGSRIAALKVI"

ORIGIN
Query Match 25.1%; Score 782; DB 6; Length 1292;
```

```
Best Local Similarity 100.0%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2325 TAAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCAATATCTAGCTTCTCG 2384
Db 486 TAAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCAATATCTAGCTTCTCG 545
Qy 2385 ATATGATTTTCATTGACTCGTGGGATCCAGGCTGGGATCCAGGCTGGTCTTATGGAGGATACCT 2444
Db 546 ATATGATTTTCATTGACTCGTGGGATCCAGGCTGGGATCCAGGCTGGTCTTATGGAGGATACCT 605
Qy 2445 CTCCTGTATGGCATTAAATGAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCC 2504
Db 606 CTCCTGTATGGCATTAAATGAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCC 665
Qy 2505 AGTCACTCTGTGGATCTTCTATGATCAGGATACAGGAAACGTTATATGGGTCAACCTGA 2564
Db 666 AGTCACTCTGTGGATCTTCTATGATCAGGATACAGGAAACGTTATATGGGTCAACCTGA 725
Qy 2565 CCAGATCAACAGGCTTATTACTTGGCCATCGGCAATGCAAGCAGAGAAAGTTCCCTC 2624
Db 726 CCAGATCAACAGGCTTATTACTTGGCCATCGGCAATGCAAGCAGAGAAAGTTCCCTC 785
Qy 2625 TGAACCAATCGTTTACTGCTCTTACATGTTTCTGGATGAGAAATGCTCAATTTTGACA 2684
Db 786 TGAACCAATCGTTTACTGCTCTTACATGTTTCTGGATGAGAAATGCTCAATTTTGACA 845
Qy 2685 TACCAGTATATTACTGAGTTTTTTAGTGGGCTGGAAGCCATATGATTACAGATCTA 2744
Db 846 TACCAGTATATTACTGAGTTTTTTAGTGGGCTGGAAGCCATATGATTACAGATCTA 905
Qy 2745 TCCTCAGGAGACACAGATAGAGTTCTCGAATCGGAGAAACATTTATGAATGCACTCT 2804
Db 906 TCCTCAGGAGACACAGATAGAGTTCTCGAATCGGAGAAACATTTATGAATGCACTCT 965
Qy 2805 TTTGCACTACTTCAAGAAAACCTTTGGATCAGTATTCCTGCTCTCAAAAGTGATATAAT 2864
Db 966 TTTGCACTACTTCAAGAAAACCTTTGGATCAGTATTCCTGCTCTCAAAAGTGATATAAT 1025
Qy 2865 TTGACCTGTGTGAACCTCTCTGTATACATCTGGCTATTAAACCAAAATGAGAGGTTAAT 2924
Db 1026 TTGACCTGTGTGAACCTCTCTGTATACATCTGGCTATTAAACCAAAATGAGAGGTTAAT 1085
Qy 2925 CAACAGAAAACACAGAAATTTGATCATCATTTTATACCTGCCATGTAACTACTCTCT 2984
Db 1086 CAACAGAAAACACAGAAATTTGATCATCATTTTATACCTGCCATGTAACTACTCTCT 1145
Qy 2985 GAAAATAAATGTGGTCCCATGCAGGGGTCTACGGTTTGTGTAGTAACTCTAATACCTTAA 3044
Db 1146 GAAAATAAATGTGGTCCCATGCAGGGGTCTACGGTTTGTGTAGTAACTCTAATACCTTAA 1205
Qy 3045 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCATAGAA 3104
Db 1206 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCATAGAA 1265
Qy 3105 TT 3106
Db 1266 TT 1267

RESULT 5
BM557438
LOCUS AGENCOURT 6578992 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466759
DEFINITION 5', mRNA sequence.
ACCESSION BM557438
VERSION BM557438.1 GI:18799430
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
```

REFERENCE 1. (bases 1 to 1042)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>;
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
Plate: LICM1967 row: k column: 16
High quality sequence stop: 697.
Location/Qualifiers
1..1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:5466759"
/tissue types="amelanotic melanoma, cell line"
/lab host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/note="Organ: Skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 24.6%; Score 769; DB 3; Length 1042;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 819; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1723 GAAGTTCCTGGCCGCGATGATCTAATATCCAGTTGATGAGTCAAGAGCTGGTATAT 1782
Db 1 GAAGTTCCTGGCCGCGATGATCTAATATCCAGTTGATGAGTCAAGAGCTGGTATAT 60
QY 1783 TTTGAAGGCACAAAGACTCCCTTTAGAGCATCACTGTAGCTAGTCAAGTTAGCTAAAT 1842
Db 61 TTTGAAGGCACAAAGACTCCCTTTAGAGCATCACTGTAGCTAGTCAAGTTAGCTAAAT 120
QY 1843 CTTGAGAGGTGACAGAGCTGACTACCGTGGCTACTCACTTCTTGTGATCAGTCAG 1902
Db 121 CTTGAGAGGTGACAGAGCTGACTACCGTGGCTACTCACTTCTTGTGATCAGTCAG 180
QY 1903 CACTGTGACTTCTTTATAAGTATAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTT 1962
Db 181 CACTGTGACTTCTTTATAAGTATAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTT 240
QY 1963 TACAAGCTATCAAGTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTTGGGCCAC 2022
Db 241 TACAAGCTATCAAGTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTTGGGCCAC 300
QY 2023 ATTTTGGATTACAGAGGTCTCTTCTGACTATACCTCCAGAAATTTCTCTTTTGA 2082
Db 301 ATTTTGGATTACAGAGGTCTCTTCTGACTATACCTCCAGAAATTTCTCTTTTGA 360
QY 2083 AGTACTACTGGATTACATTTGATGGGATGCTTACAAGCCTCATGATCAAGCCTGGA 2142
Db 361 AGTACTACTGGATTACATTTGATGGGATGCTTACAAGCCTCATGATCAAGCCTGGA 420
QY 2143 AAGAAATATCTACTGTGCTTATATATGTTGGTGGTCCCTCAGTGGTGGTGAATAT 2202
Db 421 AAGAAATATCTACTGTGCTTATATATGTTGGTGGTCCCTCAGTGGTGGTGAATAT 480
QY 2203 CGGTTTAAAGAGGTCAAGATTTTCGCTTGAATACCTAGACCTCTCTAGGTTATGTGTT 2262
Db 481 CGGTTTAAAGAGGTCAAGATTTTCGCTTGAATACCTAGACCTCTCTAGGTTATGTGTT 540

QY 2263 GTACTGATAGACACAGGGGATCTCTGTACCGAGGCTTAAATTTGAAGGCGCCTTTAAA 2322
Db 541 GTAGTGATAGACACAGGGGATCTCTGTACCGAGGCTTAAATTTGAAGGCGCCTTTANA 600
QY 2323 TATAAATGGGTCAAATAGAAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 2382
Db 601 TATAAATGGGTCAAATAGAAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 660
QY 2383 CGATATGATTTCAATTTGACTTTAGATCGTGGGATCCAGGCTGCTCTATCGAGGATAC 2442
Db 661 CGATATGATTTCAATTTGACTTTAGATCGTGGGATCCAGGCTGCTCTATCGAGGATAC 720
QY 2443 CTCTCCCTGATGGCATTAAATGCAGAGGTGAGATATCTTCAGGGTGTCTATTCTGGGCCC 2502
Db 721 CTCTCCCTGATGGCATTAAATGCAGAGGTGAGATATCTTCAGGGTGTCTATTCTGGGCCC 780
QY 2503 CCACTCACTCTGTGGATCTTCTATGATACAGGATACACGG 2542
Db 781 CCACTCACTCTGTGGATCTTCTATGATACAGGATACACGG 820
CR985650 884 bp mRNA linear EST 28-JUN-2005
LOCUS CR985650 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016N202 5',
DEFINITION mRNA sequence.
ACCESSION CR985650
VERSION CR985650.1 GI:68279535
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1. (bases 1 to 884)
AUTHORS Heil, O., Ebert, L., Hennig, S., Henze, S., Radelof, U., Schneider, D.
and Korn, B.
TITLE Human T-lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD: RZPDp9016N202.
RZPDLIB; (Human T-lymphocytes) RZPD LIB No.9016
<http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016> Contact:
Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD:
<http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=RZPDp9016N202>
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4, Primer sequence: CGGATAACAATTCCACAG.
FEATURES
Location/Qualifiers
1..884
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016N202"
/tissue type="T-Lymphocytes"
/dev stage="adult"
/lab host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
NotI; vector:
; 1st strand cDNA was prepared from mRNA obtained from
human T-Lymphocytes with a NotI - oligo(dT) primer [5',
GACTAGTCTTAGATCGCAGCGCGCCCTTTTCTTTTCTTTT 3'].

Double-stranded cDNA was ligated to Sali adaptors,
digested with NotI and cloned into the NotI and Sali sites
of the pOE80LSN_cloned vector"

ORIGIN	Query Match Best Local Similarity 24.2%; Score 755; DB 8; Length 884; Matches 805; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	ORIGIN	KEYWORDS SOURCE ORGANISM	REFERENCE AUTHORS TITLE JOURNAL COMMENT	FEATURES source
QY 25	GGCCGCTGCTACTGCGCGCTGCTTCTTAGTCGCGGTTCGCGCTGGTGTGTCACCG 84	QY 2193	EST. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 910) NIH-MGC http://mgi.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgaops-remail.nih.gov Tissue Procurement: ARCC CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) Cloned by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCMI430 row: e column: 03 High quality sequence stop: 837. Location/Qualifiers 1..910 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4649834" /tissue_type="choriocarcinoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_21" /note="Organ: placenta; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."		
DB 1	GGCGCTGCTACTGCGCGCTGCTTCTTAGTCGCGGTTCGCGCTGGTGTGTCACCG 60	DB 1			
QY 85	GGCCGCGCGCGAGGAGCCACTGCAACACGAGCCGAGTGGAGCGCGCAGCATGAAG 144	QY 2253			
DB 61	GGCGCGCGCGGAGGAGCCACTGCAACACGAGCCGAGTGGAGCGCGCAGCATGAAG 120	DB 61			
QY 145	CGGCGAGCGCGCTCCATAGCGCATGCGGACCGGTCCGGCGCGGGAAGGA 204	QY 2313			
DB 121	CGGCGAGCGCGCTCCATAGCGCATGCGGACCGGTCCGGCGCGGGAAGGA 180	DB 121			
QY 205	AAATGCAACATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTGGATATTTGAACT 264	QY 2373			
DB 181	AAATGCAACATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTGGATATTTGAACT 240	DB 181			
QY 265	GCGGACTGTGAGGAGATATTTGAATCAGGATCGGCTTAAATTTGGAGCCTTTTATGTT 324	QY 2433			
DB 241	GCAGACTGTGAGGAGATATTTGAATCAGGATCGGCTTAAATTTGGAGCCTTTTATGTT 300	DB 241			
QY 325	GAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACCGAAATATCATGCG 384	QY 2493			
DB 301	GAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACCGAAATATCATGCG 360	DB 301			
QY 385	TACATGATGGCTAAGGCCACCACATGATTTTCATGTTTGTGAAGAGGATGTCAGATGGA 444	QY 2553			
DB 361	TACATGATGGCTAAGGCCACCACATGATTTTCATGTTTGTGAAGAGGATGTCAGATGGA 420	DB 361			
QY 445	CCTCATTCAGACAGAACTTATTACCTTGCATGCTGCTGGTGAGAACAGAGAAATACACTG 504	QY 2613			
DB 421	CCTCATTCAGACAGAACTTATTACCTTGCATGCTGCTGGTGAGAACAGAGAAATACACTG 480				
QY 505	TTTTATTCTGAAATTTCCCAAACTATCAATAGAGCAGCTCTTAAATGCTCTCTTGGAG 564				
DB 481	TTTTATTCTGAAATTTCCCAAACTATCAATAGAGCAGCTCTTAAATGCTCTCTTGGAG 540				
QY 565	CCTCTTTTGGATCTTTTTCAGGCAACACTGCACTATGGAATGATTTCTGAGAGAAGAA 624				
DB 541	CCTCTTTTGGATCTTTTTCAGGCAACACTGCACTATGGAATGATTTCTGAGAGAAGAA 600				
QY 625	CTATTAGAGAAAGAAACGATTTGGAACAGTCGGAATTTGCTTTCAGATTTATCAACAA 684				
DB 601	CTATTAGAGAAAGAAACGATTTGGAACAGTCGGAATTTGCTTTCAGATTTATCAACAA 660				
QY 685	GGAAGTGAAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCTGATATTCACAA 744				
DB 661	GGAAGTGAAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCTGATATTCACAA 720				
QY 745	CCACAGGATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAATCTAGTGTCCCAAC 804				
DB 721	CCACAGGATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAATCTAGTGTCCCAAC 780				
QY 805	ATACGATGATCCCAATATGCCCC 830				
DB 781	ATACGATGATCCCAATATGCCCC 806				
RESULT 7					
LOCUS	BG479035				
DEFINITION	602526111F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4649834 5',				
ACCESSION	mRNA sequence.				
VERSION	BG479035				
	BG479035.1 GI:13411314				

COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2407 row: a column: 22 High quality sequence stop: 630.	
FEATURES		Location/Qualifiers 1..968 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6255645" /tissue_types="epidermoid carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 102" /notes="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
ORIGIN		Query Match 23.2%; Score 724; DB 3; Length 968; Best Local Similarity 99.9%; Pred. No. 0; Matches 774; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	1634	AATATAACGATCCAGTGTGGCTGCTGCTCCAGTGATTTCAAGTCTCCATCAAG 1693	
Db	1	AATATAACGATCCAGTGTGGCTGCTGCTCCAGTGATTTCAAGTCTCCATCAAG 60	
Qy	1694	AGGAGATAGCAATACCACTGCTGATCGGAAGTCTTGGCCGATGATCAATATCC 1753	
Db	61	AGGAGATAGCAATACCACTGCTGATCGGAAGTCTTGGCCGATGATCAATATCC 120	
Qy	1754	AAGTTCATGATCAGAGGCTGGTATATTTTGAAGGCACAAAGATCCCTTTAGAGC 1813	
Db	121	AAGTTGATGATCAGAGGCTGGTATATTTTGAAGGCACAAAGATCCCTTTAGAGC 180	
Qy	1814	ATCACTGTAGTACGTACGTACGTAATCTCGAGAGGTGACAAGGCTGACTGCCGTG 1873	
Db	181	ATCACTGTAGTACGTACGTACGTAATCTCGAGAGGTGACAAGGCTGACTGCCGTG 240	
Qy	1874	GCTACTCAATCTTCTGCTGATCAGTACGACTGTGACTCTTTTATAAGTAAAGTATGTA 1933	
Db	241	GCTACTCAATCTTCTGCTGATCAGTACGACTGTGACTCTTTTATAAGTAAAGTATGTA 300	
Qy	1934	ACCAGAAGATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCTCGAAGATGACCCAA 1993	
Db	301	ACCAGAAGATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCTCGAAGATGACCCAA 360	
Qy	1994	CTTGTCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTGAGTCTCTTCTGACT 2053	
Db	361	CTTGTCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTGAGTCTCTTCTGACT 420	
Qy	2054	ATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGC 2113	
Db	421	ATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGC 480	
Qy	2114	TCTACAAGCTCATGATCTACAGCTCGAAGAAATATCTCTGCTGTGTTTCATATATG 2173	
Db	481	TCTACAAGCTCATGATCTACAGCTCGAAGAAATATCTCTGCTGTGTTTCATATATG 540	
Qy	2174	GTGGTCTCAGGTGAGTGGTGAATATCGGTTTAAAGAGTCAAGTATTCGGCTGA 2233	
Db	541	GTGGTCTCAGGTGAGTGGTGAATATCGGTTTAAAGAGTCAAGTATTCGGCTGA 600	
Qy		2234 ATACCTAGCTCTCTAGGTTATGTGGTTGTAGTAGACAAACAGGGGATCCTGTCC 2293	
Db		601 ATACCTAGCTCTCTAGGTTATGTGGTTGTAGTAGACAAACAGGGGATCCTGTCC 660	
Qy		2294 GAGGGCTTAAATTTGAAGCGCTTTAAATATAAATGGGTCAATAGAAATGACGATC 2353	
Db		661 GAGGGCTTAAATTTGAAGCGCTTTAAATATAAATGGGTCAATAGAAATGACGATC 720	
Qy		2354 AGGTGAAGGACTCCCAATATCTAGCTTCGATATGATTTCAATGACTTAGATCG 2408	
Db		721 AGGTGAAGGACTCCCAATATCTAGCTTCGATATGATTTCAATGACTTAGATCG 775	
RESULT 10		BO675260 910 bp mRNA linear EST 15-JUL-2002	
LOCUS		AGENCOURT_8354972 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6275478	
DEFINITION		5' mRNA sequence.	
ACCESSION		BO675260	
VERSION		BO675260.1 GI:21786094	
KEYWORDS		EST.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		1 (bases 1 to 910)	
AUTHORS		NIH-MGC http://mgc.nci.nih.gov/.	
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL		Unpublished (1999)	
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2458 row: 1 column: 07 High quality sequence stop: 618. Location/Qualifiers 1..910 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6275478" /tissue_types="epidermoid carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 102" /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
FEATURES		Query Match 23.2%; Score 723; DB 3; Length 910; Best Local Similarity 99.8%; Pred. No. 0; Matches 823; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Qy	1635	ATATAACGATCCAGTGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAGA 1694	
Db	1	ATATAACGATCCAGTGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAGA 60	
Qy	1695	GGAGATAGCAATTAACAGTGGTGAATGGGAAGTCTTGGCCGCGCATGATCTAATATCA 1754	
Db	61	GGAGATAGCAATTAACAGTGGTGAATGGGAAGTCTTGGCCGCGCATGATCTAATATCA 120	

LOCUS
 DEFINITION DKFp43400723_r1 434 (synonym: htes3) Homo sapiens cDNA clone
 ACCESSION DKFp43400723 5', mRNA sequence.
 VERSION AL043338.1 GI:5422728
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 735)
 Blum, H., Bauerbach, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 EST (Blum, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: MIPS
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by LMU (Ludwig Maximilians University,
 Munich/Germany) within the CDNA sequencing consortium of the German
 Genome Project.
 No sl sequence available.
 This clone (DKFp43400723) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 FEATURES
 source
 1..735
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFp43400723"
 /issue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="434 (synonym: htes3)"
 /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match
 Best Local Similarity 22.7%; Score 708; DB 1; Length 735;
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2274 CAACAGGGGATCCTGTGTCACCGAGGGCTTAAATTGAGGCGCTTTAAATATTAATATAAATGGG 2333
 Db 1 CAACAGGGGATCCTGTGTCACCGAGGGCTTAAATTGAGGCGCTTTAAATATTAATATAAATGGG 60
 QY 2334 TCAATAGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCGTGATATGATTT 2393
 Db 61 TCAATAGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCGTGATATGATTT 120
 QY 2394 CATTGACTTAGATCGTGTGGGCATCCACGGCTGTCCTATGAGGATACCTCTCCCTGAT 2453
 Db 121 CATTGACTTAGATCGTGTGGGCATCCACGGCTGTCCTATGAGGATACCTCTCCCTGAT 180
 QY 2454 GGCAATTAATGCAGAGTTCAGATATCTTCAGGGTGTCTATTCCTGGGGCCCACTCACTCT 2513
 Db 181 GGCAATTAATGCAGAGTTCAGATATCTTCAGGGTGTCTATTCCTGGGGCCCACTCACTCT 240
 QY 2514 GTGGATCTTCTATGATACAGGATACACCGAGGTTATATGGTCACTCGACAGAAATCA 2573
 Db 241 GTGGATCTTCTATGATACAGGATACACCGAGGTTATATGGTCACTCGACAGAAATCA 300
 QY 2574 ACAGGGCTATTACTTAGGATCTGGCCATCCAGCAGAAAGTTCCTCTGAAACCAA 2633
 Db 301 ACAGGGCTATTACTTAGGATCTGGCCATCCAGCAGAAAGTTCCTCTGAAACCAA 360
 QY 2634 TCGTTTACTGCTCTTACATGTTTCTGGATGAGATGTCCATTTTGCACATCAGATAT 2693
 Db 361 TCGTTTACTGCTCTTACATGTTTCTGGATGAGATGTCCATTTTGCACATCAGATAT 420

QY 2694 ATTACTGAGTTTTTTAGTGAGGCTCGGAAGCCATATGATTTACAGATCTATCTCTCAGGA 2753
 Db 421 ATTACTGAGTTTTTTAGTGAGGCTCGGAAGCCATATGATTTACAGATCTATCTCTCAGGA 480
 QY 2754 GAGACACAGCATAGAGTTTCCTGAAATCGGAGCAATATGAACTGCAATCTTTTGCACCTA 2813
 Db 481 GAGACACAGCATAGAGTTTCCTGAAATCGGAGCAATATGAACTGCAATCTTTTGCACCTA 540
 QY 2814 CTTTCAAGAAACCTTGATCAGGTATTTGCTGCTCTTAAAGTATATAATTTTGAACCTGT 2873
 Db 541 CTTTCAAGAAACCTTGATCAGGTATTTGCTGCTCTTAAAGTATATAATTTTGAACCTGT 600
 QY 2874 GTAGAACTCTCTGATATACATCGGCTATTTAAACAAATGAGGAGTTTAAATCAACAGAAA 2933
 Db 601 GTAGAACTCTCTGATATACATCGGCTATTTAAACAAATGAGGAGTTTAAATCAACAGAAA 660
 QY 2934 ACACAGAAATGATCATCATTTTGTATACCTGCCATCTAATCAATCTACT 2981
 Db 661 ACACAGAAATGATCATCATTTTGTATACCTGCCATCTAATCAATCTACT 708
 RESULT 13
 LOCUS AL040398
 DEFINITION DKFp434A0714_r1 434 (synonym: htes3) Homo sapiens cDNA clone
 ACCESSION DKFp434A0714 5', mRNA sequence.
 VERSION AL040398
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 753)
 Koehrer, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 EST (Koehrer, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: MIPS
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by BMFZ (Biomedical Research Center at the Charite,
 Berlin/Germany) within the CDNA sequencing consortium of the German
 Genome Project.
 No sl sequence available.
 This clone (DKFp434A0714) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 FEATURES
 source
 1..753
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFp434A0714"
 /issue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="434 (synonym: htes3)"
 /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
 ORIGIN
 Query Match 22.4%; Score 700; DB 1; Length 753;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 750; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1609 ATTACATCTATTAAAGGAAAGCAATATAAATCAATCCAGTGGGCTGCTCTCCA 1668
 Db 1 ATTACATCTATTAAAGGAAAGCAATATAAATCAATCCAGTGGGCTGCTCTCCA 60
 QY 1669 AGTGATTTCAAGTCTCCTATCAAGAGGAGATAGCAATACCAGTGGTGAATGGGAAGT 1728

Db 541 AGAAGAACTATTAAAGAGAAAGAAACGATTTGGAACAGTCCGGAATTTGCTTTCTAGGATTA 600
 QY 678 TCACCAAGAAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTTAAAGA 737
 Db 601 TCACCAAGAAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTTAAAGA 660
 QY 738 TGGAGGGCCCAAGAGATTATACGCAACAACTTTTAGGCCCATCTAGTGGAAACTAGTTG 797
 Db 661 TGGAGGGCCCAAGAGATTATACGCAACAACTTTTAGGCCCATCTAGTGGAAACTAGTTG 720
 QY 798 TCCCAACATACGGATGATCC 818
 Db 721 TCCCAACATACGGATGATCC 741

RESULT 15
 DA571902 830 bp mRNA linear EST 07-NOV-2005
 LOCUS DA571902 HEMBB1 Homo sapiens cDNA clone HEMBB1000150 5', mRNA
 DEFINITION sequence.
 DA571902
 VERSION DA571902.1 GI:81120933
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Teuricani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushiya, N., Yonekawa, K., Ohtsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Tanase, T., Nagai, K., Ishibashi, T., Takahashi-Fujii, A., and Sugano, S. Diversification of Transcriptional Modulation: Large-scale Identification of Promoters of Human Genes
 Promoters of Human Genes
 Genome Res. 16 (1), 55-65 (2006)
 16344560
 Contact: Takao Isogai
 FLJ Project (HRI team)
 Helix Research Institute
 2-6-7 Kasuga
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com
 HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute (HRI); cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and HRI.

TITLE
 JOURNAL Promoters of Human Genes
 PUBMED 16344560
 COMMENT

FEATURES
 source Location/Qualifiers
 1..830
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HEMBB1000150"
 /tissue_type="whole embryo, mainly body"
 /dev_stage="embryo, 10 weeks"
 /clone_lib="HEMBB1"
 /note="Vector: pME18SFL3"

ORIGIN
 Query Match 22.1%; Score 689; DB 9; Length 830;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 739; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGTGCTAAGCCTCCGAGGCAAGCGCGTCTACTGCGCGCGCTCTTACTGCGCG 61
 Db 1 AGTGCTAAGCCTCCGAGGCAAGCGCGTCTACTGCGCGCGCTCTTACTGCGCG 60
 QY 62 GTTCGCCGCTGGGTTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 121

Db 61 GTTCGCCGCTGGGTTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 QY 122 AGTCGAGCG 181
 Db 121 AGTCGAGCG 180
 QY 182 CCGGCG 241
 Db 181 CCGGCG 240
 QY 242 TGGGTTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATATTGAAATGAGGAG 301
 Db 241 TGGGTTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATATTGAAATGAGGAG 300
 QY 302 CTAAATGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCACTTAAAGCTCTCTTG 361
 Db 301 CTAAATGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCACTTAAAGCTCTCTTG 360
 QY 362 CCGATACCAAGAAATATCATGGCTACATGATGGTAAAGCCACCATGATTTTCATGTTG 421
 Db 361 CCGATACCAAGAAATATCATGGCTACATGATGGTAAAGCCACCATGATTTTCATGTTG 420
 QY 422 TGAAGAGGAATGATCCAGATGGACCTCATTCAGAGAGAAATATTGAAATGAGGAG 481
 Db 421 TGAAGAGGAATGATCCAGATGGACCTCATTCAGAGAGAAATATTGAAATGAGGAG 480
 QY 482 GTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTCGAAATTCATATAGAGAG 541
 Db 481 GTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTCGAAATTCATATAGAGAG 540
 QY 542 CAGTCTTAATGCTCTCTTGGAGCCCTCTTTGGAGTCTTTTTCAGGCAACATCGGACTATG 601
 Db 541 CAGTCTTAATGCTCTCTTGGAGCCCTCTTTGGAGTCTTTTTCAGGCAACATCGGACTATG 600
 QY 602 GAATGTATTTCTCAGAGAGAAAGAACTATTAAAGAGAAAGAAAGCGATTTGAAACAGTCGAA 661
 Db 601 GAATGTATTTCTCAGAGAGAAAGAACTATTAAAGAGAAAGAAAGCGATTTGAAACAGTCGAA 660
 QY 662 TTGCTTCTTACGATTAATACCAAGGAGTGGAACTTTCTGTTTCAAGCGCGTAGTGGAA 721
 Db 661 TTGCTTCTTACGATTAATACCAAGGAGTGGAACTTTCTGTTTCAAGCGCGTAGTGGAA 720
 QY 722 TTTATCACTGTAAGATGGA 741
 Db 721 TTTATCACTGTAAGATGGA 740

RESULT 16
 CN427229 742 bp mRNA linear EST 16-MAY-2004
 LOCUS CN427229 17000600259975 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.
 DEFINITION CN427229
 ACCESSION CN427229.1 GI:47414823
 VERSION
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Brandenberger, R., Mei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W. Transcriptional characterization elucidates signaling networks that control human ES cell growth and differentiation
 Nat. Biotechnol. 22 (6), 707-716 (2004)
 15146197
 Contact: Brandenberger R
 Regenerative Medicine
 Genon Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760

TITLE
 JOURNAL Transcriptional characterization elucidates signaling networks that control human ES cell growth and differentiation
 PUBMED 15146197
 COMMENT

	Email: rbrandenberger@geron.com		
	Insert Length: 742	Std Error: 0.00.	
FEATURES	Location/Qualifiers		
source	1. .742		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue type="embryonic stem cell, retinoic acid and		
	mitogen-treated hES cell line H7"		
	/clone_lib="GRN PRENEU"		
	/notes="oligo dt' primed, full-length enriched cDNA library		
	from hES cell line H7 (p29) maintained in feeder-free		
	conditions. Embryoid bodies were generated in the presence		
	of all-trans retinoic and mitogens."		
ORIGIN	22.0%	Score 687;	DB 8; Length 742;

Query Match	22.0%	Score 687	DB 8	Length 7427	
Best Local Similarity	99.9%	Pred. No. 0			
Matches	737	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	552	GCTCTCTTGGAGGCTCTTTTGATCTTTTCAGGCAACACTGAGCTATGGAATGATTC	611		
Db	5	GCTCTCTTGGAGGCTCTTTTGATCTTTTCAGGCAACACTGAGCTATGGAATGATTC	64		
QY	612	TCGAGAGAGAACTATTAGAGAGAGAGAAACGCAATGGAAACAGCTCGGAATGCTTCTTA	671		
Db	65	TCGAGAGAGAGAACTATTAGAGAGAGAGAGAAACGCAATGGAAACAGCTCGGAATGCTTCTTA	124		
QY	672	CGATTATCACCAAGAGTGGAACTTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGT	731		
Db	125	CGATTATCACCAAGAGTGGAACTTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGT	184		
QY	732	AAAAGATGGAGGCCCAAGGATTTAGCGAACCAACCTTTAAGGCCCAATCTAGTGGAAAC	791		
Db	185	AAAAGATGGAGGCCCAAGGATTTAGCGAACCAACCTTTAAGGCCCAATCTAGTGGAAAC	244		
QY	792	TAGTTGTCCTCAACATACGGATGGATCCAAATTTATCCCGCTGATCCAGACTGGATTC	851		
Db	245	TAGTTGTCCTCAACATACGGATGGATCCAAATTTATCCCGCTGATCCAGACTGGATTC	304		
QY	852	TTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACGAGAGAGAAAGAGACT	911		
Db	305	TTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACGAGAGAGAAAGAGACT	364		
QY	912	CACATTATGCGCAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTCCG	971		
Db	365	CACATTATGCGCAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTCCG	424		
QY	972	TACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTGGGTATTTGGTGTGTCCTCAAAAGC	1031		
Db	425	TACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTGGGTATTTGGTGTGTCCTCAAAAGC	484		
QY	1032	TGAACAACCTCCCAAGTGGTGAATAATTTAGAAATTTCTATATGAGAGAAATGATGAATC	1091		
Db	485	TGAACAACCTCCCAAGTGGTGAATAATTTAGAAATTTCTATATGAGAGAAATGATGAATC	544		
QY	1092	TGAGGTGGAATTTATTCATGTTACATCCCTATGTTGGAAACAAGAGGGGCGAGATTCAT	1151		
Db	545	TGAGGTGGAATTTATTCATGTTACATCCCTATGTTGGAAACAAGAGGGGCGAGATTCAT	604		
QY	1152	CCGTTATCCTAAAAACAGTACAGCAATCCTAAAGTCATTTTAAAGATGTCAGAAATAAT	1211		
Db	605	CCGTTATCCTAAAAACAGTACAGCAATCCTAAAGTCATTTTAAAGATGTCAGAAATAAT	664		
QY	1212	GATTGATGCTGAAGAGGATCATAGATGTCATAGATAGGAACTTAATTCACCTTTTGA	1271		
Db	665	GATTGATGCTGAAGAGGATCATAGATGTCATAGATAGGAACTTAATTCACCTTTTGA	724		
QY	1272	GATTCATTTTGAAGAGT	1289		
Db	725	GATTCATTTTGAAGAGT	742		

914 bp mRNA linear EST 28-APR-2004
 BX390898 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 BX390898 Homo sapiens cDNA clone CSDDL005D02 5-PRIME, mRNA sequence.
 BX390898
 BX390898.2 GI:46844267
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 914)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 13, 2003 this sequence version replaced gi:30607432.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
 Email: seqref@genoscope.cns.fr web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by life technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7542.i
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna/s-cs08BAG053ZD07_CS05044.1&c=7542.i

	Location/Qualifiers
FEATURES	1..914
source	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="CS0DL005YDQ2"
	/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
	/cell_line="RAMOS CELL LINE"
	/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
	/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN	
	22.0%; Score 686; DB 4; Length 914;
Query Match	Best Local Similarity 99.9%; Pred.No. 0;
Matches 736;	Conservative 1; Indels 0; Gaps 0;
QY	739 GGAGGGCCACAGGATTTTACGCACAACCTTTTAAGGCCCAATCTAGTGAAACTAGTTGT 798
Db	73 GGAGGGCCACAGGATTTTACGCACAACCTTTTAAGGCCCAATCTAGTGAAACTAGTTGT 132
QY	799 CCCAACATACGGATGGATCCAAAATTATGCCCGCTGATCCAGACTGGAATTCCTTTATA 858
Db	133 CCCAACATACGGATGGATCCAAAATTATGCCCGCTGCTGATCCAGACTGGATTCGTTTTATA 192
QY	859 CATGACACCATATTTTGGATATCTAACATCGTAAACCAGAGAAGAAAGGAGACTCATTAT 918
Db	193 CATGACACCATATTTTGGATATCTAACATCGTAAACCAGAGAAGAAAGGAGACTCATTAT 252
QY	919 GTGCACATAGCTAGTCCACATCGGAAGAAGATGCCAGATCAGCTGAGTCGCTACCTTT 978
Db	253 GTGCACATAGCTAGTCCACATCGGAAGAAGATGCCAGATCAGCTGAGTCGCTACCTTT 312
QY	979 GTTCTCCAAGAAGAAATTTGATAGATATTCGGCTATTTCGGTGTCTCAAAGCTGAAACA 1038
Db	313 GTTCTCCAAGAAGAAATTTGATAGATATTCGGCTATTTCGGTGTCTCAAAGCTGAAACA 372
QY	1039 ACTCCCAGTGGTGGTAAAAATTCCTTAGAATTCCTATGAGAAAATGATGAATCTGAGGTG 1098

373 ACTCCAGTGGTGTAAATTTCTTAGAATCTTATGAGAAAAATGATGAATCTGAGTG 432
 1099 GAATATTTATGTTTACATCCCTATGTTGAAACAAGGAGGCGAGATTCATTCGGTTAT 1158
 433 GAATATTTATGTTTACATCCCTATGTTGAAACAAGGAGGCGAGATTCATTCGGTTAT 492
 1159 CCTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGTGTGAGAAATATGATTCAT 1218
 493 CCTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGTGTGAGAAATATGATTCAT 552
 1219 GCTGAAGGAAGATCATAGATGTCATAGATAAGGAATTAATTCACCTTTTGAATTTCTA 1278
 553 GCTGAAGGAAGATCATAGATGTCATAGATAAGGAATTAATTCACCTTTTGAATTTCTA 612
 1279 TTTGAAGGAGTTGAATATATGTCAGAGCTGGATGAGTCCCTGAGGGAATATGCTTGG 1338
 613 TTTGAAGGAGTTGAATATATGTCAGAGCTGGATGAGTCCCTGAGGGAATATGCTTGG 672
 1339 TCCATCTCTAGATGCTCCAGACTCCCTACAGATGAGTGGTCTCTGAGGGAATATGCTTGG 732
 673 TCCATCTCTAGATGCTCCAGACTCCCTACAGATGAGTGGTCTCTGAGGGAATATGCTTGG 732
 1399 TTTATCCAGTACAGATGATGTTATGGAAGGAGGAGCACTATGAGTGTGAGTGTGAGTGT 1458
 733 TTTATCCAGTACAGATGATGTTATGGAAGGAGGAGCACTATGAGTGTGAGTGTGAGTGT 792
 1459 TCTGTGACGCCCACTAAT 1475
 793 TCTGTGACGCCCACTAAT 809

RESULT 18
 CK000768
 LOCUS
 DEFINITION AGENCOURT 1636862 NIH_MGC.221 Homo sapiens cDNA clone
 IMAGE:30707852 5', mRNA sequence.
 CK000768
 CK000768.1 GI:38526802
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo; 224
 1 (bases) 1 (6724)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute
 Bldg 31 Rm10A07 Bethesda, MD 20892
 Email: cgep@nci.nih.gov
 Tissue Procurement: James Martin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAMI073 row: b column: 21
 High quality sequence stop: 678.
 Location/Qualifiers
 1..724
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30707852"
 /lab_host="DH10B Tona"
 /clone_lib="NIH_MGC_221"
 /note="Organ: mixed; Vector: pX-Asc; Site 1: EcoRI;
 Site 2: NotI; Library is oligo-dT primed and directionally
 cloned. Denatured RNA was size fractionated on a 1% agarose

FEATURES

source

RESULT 19
 CR994557
 LOCUS
 DEFINITION
 CR994557

CR994557 820 bp mRNA linear EST 28-JUN-2005
 CR994557 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016A0619 5',
 mRNA sequence.
 CR994557

gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to RNA size fraction, ligated with
 EcoR I adaptor digested with Not I and then cloned
 directionally into pX-ABC vector. Average insert size
 4-5kb. Adaptors 5' (ATTCGACAGG)3' and 5'd
 (CCTCGTGGC)3'. Linker sequence - GCGGCGCTGAGAGCC T18.
 Sequencing primers 3' end: T3 promoter primer 5'd
 (ATTACCTCTCACTAAGCA)3'. 5' End: T7 promoter primer 5'd
 (TAATACGCTCACTAAGG)3'. Library was constructed in the
 laboratory of M. Bento Soares. Note: this is a NIH_MGC
 Library"

ORIGIN

Query Match 22.0%; Score 685; DB 5; Length 724;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 685; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 23 AAGCGCGCTCTACTCGCCCGCTGCTTTAGTCCGCGTTCGCCCGCTGGGTGTGCAC 82
 Db 8 AAGCGCGCTCTACTCGCCCGCTGCTTTAGTCCGCGTTCGCCCGCTGGGTGTGCAC 67
 QY 83 CGCGCGCGCGCGCGAGGAGCACTGCAACAGGACCGGAGTGGAGCGCGCGAGCATGA 142
 Db 68 CGCGCGCGCGCGCGAGGAGCACTGCAACAGGACCGGAGTGGAGCGCGCGAGCATGA 127
 QY 143 AGCGCGCGAGCGCGCTCCATAGCGCACGTCGGGACGCTCCGGGCGGTCGGGCGGGAAG 202
 Db 128 AGCGCGCGAGCGCGCTCCATAGCGCACGTCGGGACGCTCCGGGCGGTCGGGCGGGAAG 187
 QY 203 GAAATGCAACATGCGAGCAATGAAACAGACAGCTGGGTGTGAGATATTGAAA 262
 Db 188 GAAATGCAACATGCGAGCAATGAAACAGACAGCTGGGTGTGAGATATTGAAA 247
 QY 263 CTGCGGACTGTGAGGAGAAATTAATTAATCAAGGACGCTTAATTTGGAGCCCTTTTATG 322
 Db 248 CTGCGGACTGTGAGGAGAAATTAATTAATCAAGGACGCTTAATTTGGAGCCCTTTTATG 307
 QY 323 TTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTCCGACATACCAAGAAATATCATG 382
 Db 308 TTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTCCGACATACCAAGAAATATCATG 367
 QY 383 GCTACATGATGGGTAAAGGACCAACATGATTTCAATTTGTGAGAGGATGATCCAGATG 442
 Db 368 GCTACATGATGGGTAAAGGACCAACATGATTTCAATTTGTGAGAGGATGATCCAGATG 427
 QY 443 GACCTCATTCAGACAGAAATCTATTACCTTGGCATCTGTGAGACAGAGAAATACAC 502
 Db 428 GACCTCATTCAGACAGAAATCTATTACCTTGGCATCTGTGAGACAGAGAAATACAC 487
 QY 503 TGTTTTATCTGAAATTTCCAAACTCAATAGAGAGAGCTCTTAATGCTCTCTTGA 562
 Db 488 TGTTTTATCTGAAATTTCCAAACTCAATAGAGAGAGCTCTTAATGCTCTCTTGA 547
 QY 563 AGCTCTCTTTGGATCTTTTCAGGCAACATGGAATGGAATGATTTCTCGAGAGAG 622
 Db 548 AGCTCTCTTTGGATCTTTTCAGGCAACATGGAATGGAATGATTTCTCGAGAGAG 607
 QY 623 AACTATTAGAGAAAGAAAGCACTTGAACAGTCGGAATGCTCTTACGATTATCACC 682
 Db 608 AACTATTAGAGAAAGAAAGCACTTGAACAGTCGGAATGCTCTTACGATTATCACC 667
 QY 683 AAGGAAGTGGACATTTCTGTTCA 707
 Db 668 AAGGAAGTGGACATTTCTGTTCA 692

ORIGIN

Query Match 21.7%; Score 676; DB 1; Length 864;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2268 GATAGACAACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGGGCCCTTTAAATATATA 2327
DB 1 GATAGACAACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGGGCCCTTTAAATATATA 60

QY 2328 AATGGGTCAAAATAGAAATGACGATCAGGTGAGGAGCTCCATATCTAGCTTCTCGATA 2387
DB 61 AATGGGTCAAAATAGAAATGACGATCAGGTGAGGAGCTCCATATCTAGCTTCTCGATA 120

QY 2388 TGATTTCAITGACTTAGATCTGTGGGCATCCAGGCTGGTCTATGGAGGATACCTCTC 2447
DB 121 TGATTTCAITGACTTAGATCTGTGGGCATCCAGGCTGGTCTATGGAGGATACCTCTC 180

QY 2448 CCTGATGGCATTAATGACAGGTCAGATATCTTCAGGTTGCTATTTGGGGCCCAAGT 2507
DB 181 CCTGATGGCATTAATGACAGGTCAGATATCTTCAGGTTGCTATTTGGGGCCCAAGT 240

QY 2508 CACTCTGTGGATCTTCTATGATACAGGATACAGGACGTTATATGGGTCAACCTGACCA 2567
DB 241 CACTCTGTGGATCTTCTATGATACAGGATACAGGACGTTATATGGGTCAACCTGACCA 300

QY 2568 GAATGAAACAGGGCTATTACTTAGGATCTGTGGGCATGCAAGCAGAAAAGTTCCCTCTGA 2627
DB 301 GAATGAAACAGGGCTATTACTTAGGATCTGTGGGCATGCAAGCAGAAAAGTTCCCTCTGA 360

QY 2628 ACCAAATCGTTTACTCTCTTACATGGTTCTCGATGAGAAATGTCCATTTTGCACATAC 2687
DB 361 ACCAAATCGTTTACTCTCTTACATGGTTCTCGATGAGAAATGTCCATTTTGCACATAC 420

QY 2688 CAGTATATTACTCAGTTTTTATGAGGGCTGGAAAGCCAPATGATTTACAGATCTATCC 2747
DB 421 CAGTATATTACTCAGTTTTTATGAGGGCTGGAAAGCCAPATGATTTACAGATCTATCC 480

QY 2748 TCAGGAGACACAGCATTAAGATTCCTGAATCGGAGAAATATGAACTGCACTCTTT 2807
DB 481 TCAGGAGACACAGCATTAAGATTCCTGAATCGGAGAAATATGAACTGCACTCTTT 540

QY 2808 GCACTACCTTCAAGAAACCTTGGATCAGTATTTGCTCTCTAAAAGTGATATAATTTG 2867
DB 541 GCACTACCTTCAAGAAACCTTGGATCAGTATTTGCTCTCTAAAAGTGATATAATTTG 600

QY 2868 ACTGTGTAGACTCTCTGGTATACATGCTATTTAACCAATGAGGAGTTTAAATCAA 2927
DB 601 ACTGTGTAGACTCTCTGGTATACATGCTATTTAACCAATGAGGAGTTTAAATCAA 660

QY 2928 CAGAAACACAGAAAT 2943
DB 661 CAGAAACACAGAAAT 676

RESULT 21
AY411616
LOCUS
DEFINITION
Pan troglodytes DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY411616
VERSION
AY411616.1
KEYWORDS
GSS
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.
1 (bases 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE

JOURNAL
COMMENT
FEATURES

Submitted (18-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
source
1..2649
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..2649
/gene="DPP8"
/locus_tag="HCM4283"

ORIGIN

Query Match 21.6%; Score 674; DB 14; Length 2649;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 824; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1750 ATCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTA 1809
DB 1537 ATCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTA 1596

QY 1810 GAGCATCACCTCTAGCTAGTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACTGAC 1869
DB 1597 GAGCATCACCTCTAGCTAGTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACTGAC 1656

QY 1870 CGTGGCTACTCACATTTCTGTCTGCATCAGTCAGCACTGTGACTTCTTTATAGTAGTAT 1929
DB 1657 CGTGGCTACTCACATTTCTGTCTGCATCAGTCAGCACTGTGACTTCTTTATAGTAGTAT 1716

QY 1930 AGTAACCAAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCTCGAAGATGAC 1989
DB 1717 AGTAACCAAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCTCGAAGATGAC 1776

QY 1990 CCACTTGCAAAAACAAAGGAATTTTGGGCCACCAATTTGGATTTCAGAGGTCCTCTTCT 2049
DB 1777 CCACTTGCAAAAACAAAGGAATTTTGGGCCACCAATTTGGATTTCAGAGGTCCTCTTCT 1836

QY 2050 GACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGG 2109
DB 1837 GACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGG 1896

QY 2110 ATGCTCTACAAGCTCTCATGATCTACAGCTCGAAAGAAATATCTTACTGTGCTTTCATA 2169
DB 1897 ATGCTCTACAAGCTCTCATGATCTACAGCTCGAAAGAAATATCTTACTGTGCTTTCATA 1956

QY 2170 TATGGTGGTCTCAGGTGAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCCG 2229
DB 1957 TATGGTGGTCTCAGGTGAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCCG 2016

QY 2230 TTGAATACCTTAGCTCTCTAGTTTCTGGTTAGTGTAGTAGATAGACACAGGGATCTGT 2289
DB 2017 TTGAATACCTTAGCTCTCTAGTTTCTGGTTAGTGTAGTAGATAGACACAGGGATCTGT 2076

QY 2290 CACCGAGGCTTAAATTTGAAGGGCCCTTTAAATATAAAATGGGTCAATAGAAATGAC 2349
DB 2077 CACCGAGGCTTAAATTTGAAGGGCCCTTTAAATATAAAATGGGTCAATAGAAATGAC 2136

QY 2350 GATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATGATTTCAATTTGACTTAGATCT 2409
DB 2137 GATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATGATTTCAATTTGACTTAGATCT 2196

QY 2410 GTGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGTCATTATATGACAGG 2469

```

Db      2197 GTGGGATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCAATTAATGACAGG 2256
      |||
Qy      2470 TCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGAT 2529
      |||
Db      2257 TCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGAT 2316
      |||
Qy      2530 ACAGGATACACGGAACGTTATATGAGGTACCTCTGACCAAGATGAACA 2576
      |||
Db      2317 ACAGGATACACGGAACGTTATATGAGGTACCTCTGACCAAGATGAACA 2363
      |||

RESULT 22
LOCUS    BE733691                1002 bp    mRNA    linear    EST 15-SEP-2000
DEFINITION 60156288F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3844101 5',
      mRNA sequence.
ACCESSION BE733691
VERSION    BE733691.1 GI:10147786
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
      Hominoidea; Homo.
REFERENCE  1 (bases 1 to 1002)
      NIH-MGC http://mgc.nci.nih.gov/
      National Institutes of Health, Mammalian Gene Collection (MGC)
      Unpublished (1999)
      Contact: Robert Strausberg, Ph.D.
      Email: cgapbs-r@mail.nih.gov
      Tissue Procurement: ATCC
      cDNA Library Preparation: Ling Hong/Rubin Laboratory
      cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
      DNA Sequencing by: Incyte Genomics, Inc.
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
      Plate: LfCM539 row: p column: 22
      High quality sequence stop: 672.
      Location/Qualifiers
        1..1002
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="IMAGE:3844101"
          /tissue_type="choriocarcinoma"
          /lab_host="DH10B (phage-resistant)"
          /clone_lib="NIH MGC 21"
          /notes="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
          Site 2: EcoRI; cDNA made by oligo-dT priming.
          Directionally cloned into EcoRI/XhoI sites using the
          following 5' adaptor: GGCACGAG(G). Size-selected >500bp
          for average insert size 1.8kb. Library constructed by
          Ling Hong in the laboratory of Gerald M. Rubin (University
          of California, Berkeley) using ZAP-cDNA synthesis kit
          (Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

```

Query Match      21.5%; Score 672; DB 7; Length 1002;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2226 CCCTTTGAATACCTAGCTCTCTAGGTTATGTGTTGTAGTGATGACACACAGGGGATC 2285
      |||
Db      15  CCCTTTGAATACCTAGCTCTCTAGGTTATGTGTTGTAGTGATGACACACAGGGGATC 74
      |||

Qy      2286 CTCTCCACCGGGCTTAATTTGAAGGGCCCTTTAAATATATAAATGGTCAATAGAAAT 2345
      |||
Db      75  CTGTCCACCGGGCTTAATTTGAAGGGCCCTTTAAATATATAAATGGTCAATAGAAAT 134
      |||

Qy      2346 TGACGATCAGGTGGAAGGACTCCAATATCTAGCTCTTCGATATGATTTCTTACGACTTAGA 2405
      |||
Db      135 TGACGATCAGGTGGAAGGACTCCAATATCTAGCTCTTCGATATGATTTCTTACGACTTAGA 194
      |||
```

```

Qy      2406 TCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAATTAATGCA 2465
      |||
Db      195  TCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAATTAATGCA 254
      |||
Qy      2466 GAGTTCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTA 2525
      |||
Db      255  GAGTTCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTA 314
      |||
Qy      2526 TGATACAGGATACACGGAACGTTATATGAGGTACCTCTGACCAAGATGAACAGGGCTATTA 2585
      |||
Db      315  TGATACAGGATACACGGAACGTTATATGAGGTACCTCTGACCAAGATGAACAGGGCTATTA 374
      |||
Qy      2586 CTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTGCAACCAATCGTTTACTGCT 2645
      |||
Db      375  CTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTGCAACCAATCGTTTACTGCT 434
      |||
Qy      2646 CTTACATGGTTTCTCGATGAGAAATGTCATTTTGCACATACAGTATATTAATCTAGTGT 2705
      |||
Db      435  CTTACATGGTTTCTCGATGAGAAATGTCATTTTGCACATACAGTATATTAATCTAGTGT 494
      |||
Qy      2706 TTTAGTAGGGCTGGAAAGCCATATGATTACAGATCTATCTCAGGAGAGACACAGCAT 2765
      |||
Db      495  TTTAGTAGGGCTGGAAAGCCATATGATTACAGATCTATCTCAGGAGAGACACAGCAT 554
      |||
Qy      2766 AAGAGTTCTGGAATCGGAGAACATTTATGAATCTTTTGGCACTACCTTCAAGAAAA 2825
      |||
Db      555  AAGAGTTCTGGAATCGGAGAACATTTATGAATCTTTTGGCACTACCTTCAAGAAAA 614
      |||
Qy      2826 CTTGGATCACGTAATGCTGCTCTAAAAGTATATAATTTTGACCTGTGTAGAACTCTCT 2885
      |||
Db      615  CTTGGATCACGTAATGCTGCTCTAAAAGTATATAATTTTGACCTGTGTAGAACTCTCT 674
      |||
Qy      2886 GGTATACACTGG 2897
      |||
Db      675  GGTATACACTGG 686
      |||

RESULT 23
LOCUS    CX787143                671 bp    mRNA    linear    EST 02-MAR-2005
DEFINITION HSC3 81 D06.g1 A036 NIH MGC 260 Homo sapiens cDNA clone
      IMAGE:7483429 5', mRNA sequence.
ACCESSION CX787143
VERSION    CX787143.1 GI:58303933
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
      Hominoidea; Homo.
REFERENCE  1 (bases 1 to 671)
      NIH-MGC http://mgc.nci.nih.gov/
      National Institutes of Health, Mammalian Gene Collection (MGC)
      Unpublished (1999)
      Contact: Daniela S. Gerhard, Ph.D.
      Office of Cancer Genomics
      National Cancer Institute / NIH
      Bldg. 31 Rm10A07 Bethesda, MD 20892
      Email: cgapbs-r@mail.nih.gov
      Tissue Procurement: Bresagen, Inc.
      cDNA Library Preparation: Express Genomics, Inc.
      cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
      DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
      University of Georgia
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      http://image.llnl.gov
      Plate: LLAMI5806 row: g column: 11
      Seq primer: JENREV (CAGGAACAGCTATGACC)
      High quality sequence stop: 671.
      Location/Qualifiers
        1..671
```

FEATURES
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7483429"
/sex="male"
/tissue_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="NIH MGC_260"
/notes="Vector: pExpress-1; Site 1: NotI; Site 2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos. Cell line id and NIH Registry designation is BG01. Positive for SSEA3, SSEA4, Tra 1-81, CD9, Alk Phos, Oct4 and Nanog expression; negative for SSEA1 expression. Passage number 21. DNA primed using oligo-dT primer: 5'-pCAGTACTCTAGATCGAGCGGCCCGCC(7)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Note: this is a Mammalian Gene Collection library."

```

ORIGIN

```

Query Match      21.5%; Score 671; DB 9; Length 671;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1650 TGGTGGGCTGCTCTCAAGTGAATTCATCAAGTGTCTATCAAGAGGAGATAGCAATTAC 1709
Db 1 TGGTGGGCTGCTCTCAAGTGAATTCATCAAGTGTCTATCAAGAGGAGATAGCAATTAC 60

QY 1710 CAGTGTGTATGGGAAGTCTTGGCCGGATGGATCTAATATCAAGTTGATGAAGTCAG 1769
Db 61 CAGTGTGTATGGGAAGTCTTGGCCGGATGGATCTAATATCAAGTTGATGAAGTCAG 120

QY 1770 AAGCTGTGTATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACTGTACCTAGT 1829
Db 121 AAGCTGTGTATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACTGTACCTAGT 180

QY 1830 CAGTACGTAAATCTCGGAGAGGTGACCAAGGTGACTGACCGTGGCTACTCACATCTCTG 1889
Db 181 CAGTACGTAAATCTCGGAGAGGTGACCAAGGTGACTGACCGTGGCTACTCACATCTCTG 240

QY 1890 CTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAGGATCCACA 1949
Db 241 CTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAGGATCCACA 300

QY 1950 CTGTGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACCAAGAGA 2009
Db 301 CTGTGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACCAAGAGA 360

QY 2010 ATTTTGGGCCACATTTTGGATTTCAGCAGGTCTCTTCTGCTACTACTCTCCAGAAAT 2069
Db 361 ATTTTGGGCCACATTTTGGATTTCAGCAGGTCTCTTCTGCTACTACTCTCCAGAAAT 420

QY 2070 TTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGGATGCTCTACAGCTCTCATGA 2129
Db 421 TTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGGATGCTCTACAGCTCTCATGA 480

QY 2130 TCTCAGCCTGGAAAGAAATATCTCTGCTGTGTTTCATATGTTGCTCTCAGGTGCA 2189
Db 481 TCTCAGCCTGGAAAGAAATATCTCTGCTGTGTTTCATATGTTGCTCTCAGGTGCA 540

QY 2190 GTTGTGTAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCT 2249
Db 541 GTTGTGTAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCT 600

QY 2250 AGGTTATGTGTTGCTAGTATGACCAACAGGGGATCTCTGTCACCGAGGGCTTAAATTTGA 2309
Db 601 AGGTTATGTGTTGCTAGTATGACCAACAGGGGATCTCTGTCACCGAGGGCTTAAATTTGA 660

```

```

QY 2310 AGCGGCTTTA 2320
Db 661 AGCGGCTTTA 671

RESULT 24
BU687687/c
LOCUS DEFINITION
BU687687
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 804)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Source
1..804
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-adu-o-02-0-UI"
/tissue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="UI-CF-EC1"
/clone_lib="UI-CF-EC1"
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
ad118 tail. The sequence tag for this library is
AAGTGTCTTAC.
TAG ISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_LIB=UI-CF-EC1
TAG_SEQ=AAGTGTCTTAC"

ORIGIN
Query Match      21.2%; Score 663; DB 3; Length 804;

```


Db 601 TCCCTTCTGCTACTACTCTCCAGAAATTTCTCTTTTGNAGTACTACTGGATTAC 660

RESULT 26
 BX372276/c
 LOCUS
 DEFINITION BX372276 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 ACCESSION BX372276
 VERSION BX372276
 KEYWORDS BX372276.1 GI:30448117
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 910)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
 end enriched double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7542.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?e=CS0BA10432H06_CS04080_1&c=7542.r

FEATURES
 source

Location/Qualifiers
 1..910
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL005YD02"
 /cell_type="B CELLS (RAMOS CELL LINE)"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
 25-NORMALIZED"
 /notes="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match
 Best Local Similarity 20.9%; Score 652; DB 4; Length 910;
 Matches 702; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1293 ATATATGCGAGCTGGATGACTCTCTGAGGGAATATGCTTGGCTCATCTACTAGA 1352
 Db 743 ATATATGCGAGCTGGATGACTCTCTGAGGGAATATGCTTGGCTCATCTACTAGA 684

Qy 1353 TCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACTCACTGAATTTATCCAGTAGA 1412
 Db 683 TCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACTCACTGAATTTATCCAGTAGA 624

Qy 1413 AGATGATGTTATGAAAGGAGAGACTCTTGTAGTCACTGCTGATCTTGTGAGCCCACT 1472
 Db 623 AGATGATGTTATGAAAGGAGAGACTCTTGTAGTCACTGCTGATCTTGTGAGCCCACT 564

Qy 1473 AATTATCTATGAAGAAACAACAGACATCTGGTAATATCCAGTCACTTTCAAGTTT 1532
 Db 563 AATTATCTATGAAGAAACAACAGACATCTGGTAATATCCAGTCACTTTCAAGTTT 504

Qy 1533 TCCCAAGTCAAGAGAGAAATTCAGTTTATTTTGGCTCTGATGCAAAACAGTTT 1592
 Db 503 TCCCAAGTCAAGAGAGAAATTCAGTTTATTTTGGCTCTGATGCAAAACAGTTT 444

Qy 1593 CCCTCATTTATACAAATTTACATCTATTTTAAAGAAAGCAATATATAAGATCCAGTGG 1652
 Db 443 CCCTCATTTATACAAATTTACATCTATTTTAAAGAAAGCAATATATAAGATCCAGTGG 384

Qy 1653 TGGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATACCAG 1712
 Db 383 TGGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATACCAG 324

Qy 1713 TGGTGAATGGGAAGTTCTTGGCCGCGATGATCTAATATCAAGTTGTAGTCAAGTCAAG 1772
 Db 323 TGGTGAATGGGAAGTTCTTGGCCGCGATGATCTAATATCAAGTTGTAGTCAAGTCAAG 264

Qy 1773 GCTGGTATATTTTGAAGCACCACCAAGACTCCCTTTAGAGCATCACTGTAGTGTAG 1832
 Db 263 GCTGGTATATTTTGAAGCACCACCAAGACTCCCTTTAGAGCATCACTGTAGTGTAG 204

Qy 1833 TTACCTAAATCTGGAGAGGTGACCAAGGCTGACCTGACCTGCTGCTGCTGCTGCTG 1892
 Db 203 TTACCTAAATCTGGAGAGGTGACCAAGGCTGACCTGACCTGCTGCTGCTGCTGCTG 144

Qy 1893 CATCAGTCAGCTGTGACTTCTTTATTAAGTAAAGTATAGTACCAAGAGATCCCACTG 1952
 Db 143 CATCAGTCAGCTGTGACTTCTTTATTAAGTAAAGTATAGTACCAAGAGATCCCACTG 84

Qy 1953 TGTGTCCCTTTTACAAGCTATCAAGTCTCAAGATGACCCCACT 1995
 Db 83 TGTGTCCCTTTTACAAGCTATCAAGTCTCAAGATGACCCCACT 41

RESULT 27
 DA415588
 LOCUS
 DEFINITION DA415588 BRTHA3 Homo sapiens cDNA clone BRTHA3024429 5', mRNA
 ACCESSION DA415588
 VERSION DA415588
 KEYWORDS DA415588.1 GI:81137524
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 657)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Tsutsumi, K., Kishigaki, H.,
 Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
 Yoneyama, T., Otsuka, R., Kanda, K., Tokoi, T., Kondo, H., Wagatsuma, M.,
 Tanase, T., Nagai, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
 Diversification of Transcriptional Modulation: Large-scale
 Identification and Characterization of Putative Alternative
 Promoters of Human Genes
 Genome Res. 16 (1), 55-65 (2006)
 16344560
 CONTACT: Takao Isogai
 FLJ Project (HRI team)
 Helix Research Institute
 2-6-7 Kasuga, Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-ed@helix.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction;
 Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
 Center, National Institute of Technology (RAB) and Biotechnology
 pass sequencing: RAB.
 FEATURES
 source
 1..657
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

Thu Jun 22 09:04:26 2006

Email: cgaabs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10684 row: d column: 21
High quality sequence stop: 784.
Location/Qualifiers
1. 789

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4797668"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN

Query Match 20.7%; Score 647; DB 2; Length 789;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 697; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2010 ATTTGGCCACCATTTTGGATTTCAGCAGGTCCTCTCTCCTCAGTATATCTCTCCAGAAAT 2069
DB 9 ATTTGGCCACCATTTTGGATTTCAGCAGGTCCTCTCTCCTCAGTATATCTCTCCAGAAAT 68
QY 2070 TTTCTCTTTTGAAGATCTACTCGGATTTACATTTGATGGATGCTCTCAAGCCTCATGA 2129
DB 69 TTTCTCTTTTGAAGATCTACTCGGATTTACATTTGATGGATGCTCTCAAGCCTCATGA 128
QY 2130 TCTACAGCCTGGAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTCTCAGGTGCA 2189
DB 129 TCTACAGCCTGGAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTCTCAGGTGCA 188
QY 2190 GTTGTGATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCT 2249
DB 189 GTTGTGATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCT 248
QY 2250 AGGTTATGTTGTTAGTATAGACACAGGGATCTCTGTCACCGAGGCTTAAATTTGA 2309
DB 249 AGGTTATGTTGTTAGTATAGACACAGGGATCTCTGTCACCGAGGCTTAAATTTGA 308
QY 2310 AGGCGCTTTAAATATAAAATGGGTCAAAATGAGATGACATCAGTGGAGGACTCCA 2369
DB 309 AGGCGCTTTAAATATAAAATGGGTCAAAATGAGATGACATCAGTGGAGGACTCCA 368
QY 2370 ATATCTAGCTTCTCGATATGATTTCAATGACTTATAGATCGTGTGGGCAATCCAGGCTGGTC 2429
DB 369 ATATCTAGCTTCTCGATATGATTTCAATGACTTATAGATCGTGTGGGCAATCCAGGCTGGTC 428
QY 2430 CTATGGAGATATCTCTCCCTGATGGCAATTAATGAGAGTCAAGTATCTTCAGGCTTGC 2489
DB 429 CTATGGAGATATCTCTCCCTGATGGCAATTAATGAGAGTCAAGTATCTTCAGGCTTGC 488
QY 2490 TATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAACTTGA 2549
DB 489 TATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTTGA 548
QY 2550 TATGGTCCACCTTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCATCAAGC 2609

/clone="BRTHA3024429"
/tissue_type="thalamus"
/clone_lib="BRTHA3"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 20.8%; Score 649; DB 9; Length 657;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 AAAGCTCCGAGGCAAGCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGGTTGCG 67
DB 9 AAAGCTCCGAGGCAAGCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGGTTGCG 68
QY 68 CGCTCGGTTGTACCGCGCGCGCGCGCGAGGAGCACTGCAACAGGACCGGAGTGA 127
DB 69 CGCTCGGTTGTACCGCGCGCGCGCGCGAGGAGCACTGCAACAGGACCGGAGTGA 128
QY 128 GCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGGCGCACTCGGAGCGGTCGGGC 187
DB 129 GCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGGCGCACTCGGAGCGGTCGGGC 188
QY 188 GGGCGCGGGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGACAGCTGGGTG 247
DB 189 GGGCGCGGGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGACAGCTGGGTG 248
QY 248 TTGAGATATTTGAAACTCGGCACTGTGAGGAGATATTGAATCAGAGATCGGCTTAAAT 307
DB 249 TTGAGATATTTGAAACTCGGCACTGTGAGGAGATATTGAATCAGAGATCGGCTTAAAT 308
QY 308 TGGAGCCTTTTATGTTGAGCGGTATTTCTGAGGTGAGCTTTTAAAGCTGTTGCCGATA 367
DB 309 TGGAGCCTTTTATGTTGAGCGGTATTTCTGAGGTGAGCTTTTAAAGCTGTTGCCGATA 368
QY 368 CCAGAAATATCATGCTACATGAGCTTAAGGCAACCATGATTTTCATGTTTGTGAAGA 427
DB 369 CCAGAAATATCATGCTACATGAGCTTAAGGCAACCATGATTTTCATGTTTGTGAAGA 428
QY 428 GGAATGATCGAGTGGACCTTATTCAGACAGATCTATTACCTTGCCATGCTGGTGAGA 487
DB 429 GGAATGATCGAGTGGACCTTATTCAGACAGATCTATTACCTTGCCATGCTGGTGAGA 488
QY 488 ACAGAGAAATACACTGTTTATTTCTGAAATCCCAAACTATCATAGCAGCAGTCT 547
DB 489 ACAGAGAAATACACTGTTTATTTCTGAAATCCCAAACTATCATAGCAGCAGTCT 548
QY 548 TAATGCTCTTTGGAGCCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGT 607
DB 549 TAATGCTCTTTGGAGCCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGT 608
QY 608 ATTTCTGAGAGAGAGACTATTATAGAGAAAGAAAAACGATTTGGAACAGT 656
DB 609 ATTTCTGAGAGAGAGACTATTATAGAGAAAGAAAAACGATTTGGAACAGT 657

RESULT 28
LOCUS BG709118 789 bp mRNA linear EST 07-MAY-2001
DEFINITION 602675382P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4797668 5',
mRNA sequence.

ACCESSION BG709118.1 GI:13987136

VERSION EST.

KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 789)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.

```

Db      549  TATGGTCACCTGACCTGAATGAACAGGCTATTACTTAGGATCTGGCCATGCAAGC 608
QY      2610 AGAAAGTCCCTCTGAACCAATCGTTTACTGCTCTTACATGCTTCTGGATGAGAA 2669
Db      609  AGAAAGTCCCTCTGAACCAATCGTTTACTGCTCTTACATGCTTCTGGATGAGAA 668
QY      2670 TGTCCATTTTGCATACACCATATATTTACTGAGTCTTTT 2707
Db      669  TGTCCATTTTGCATACACCATATATTTACTGAGTCTTTT 706

RESULT 29
LOCUS   BG390103
DEFINITION 60241594F1 NIH_MGC_92 Homo sapiens CDNA clone IMAGE:4524323 5',
mRNA sequence.
ACCESSION BG390103
VERSION   BG390103.1 GI:13283649
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominiidae; Homo.
          1 (bases 1 to 669)
          NIH-MGC http://mgi.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strauberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue procurement: ATCC
          CDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
          Clone distribution: Incyte Genomics, Inc.
          found through the I.M.A.G.E. Consortium/LLNL at:
          Plate: L1AM10428
          High quality sequence stop: 669.
          Location/Qualifiers
            1..669
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone_image="4524323"
              /tissue_image="embryonal carcinoma, cell line"
              /lab_host="DH10B (phage-resistant)"
              /note="Organ: Testis; Vector: pCMV-SPORT6; Site 1: NotI;
              Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
              Average insert size 2.5 kb. Library enriched for
              full-length clones and constructed by Life Technologies.
              Note: this is a NIH_MGC Library."

ORIGIN
Query Match 20.6%; Score 644; DB 2; Length 669;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 644; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      881  CTACATCGTAAACAGAGAGAAAGGAGACTCACTTATGTGCACATGAGCTAGCCAAACA 940
Db      1    CTACATCGTAAACAGAGAGAAAGGAGACTCACTTATGTGCACATGAGCTAGCCAAACA 60
QY      941  TGGAGAGATGCGCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAACAAATTCATA 1000
Db      61   TGGAGAGATGCGCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAACAAATTCATA 120
QY      1001 GATATTCGTGCTATTGGTGGTGTCTCCAAAAGCTGAAACAACTCCCAAGTGGGTAAATTC 1060
Db      121  GATATTCGTGCTATTGGTGGTGTCTCCAAAAGCTGAAACAACTCCCAAGTGGGTAAATTC 180
QY      1061 TTAGATTTATTAAGAAGAAATATGATGATCTGAGGTGGAATATTTCATGTTACATCCC 1120

```

```

Db      181  TTAGATTTCTATATGAAGAAATATGATGATCTGAGGTGGAATATTATTCATGTTACATCCC 240
QY      1121  CTATGTTGGAAACAGAGGGGAGATTCATTCCTGTTTATCTTAAACAGGTACAGCAATC 1180
Db      241  CTATGTTGGAAACAGAGGGGAGATTCATTCCTGTTTATCTTAAACAGGTACAGCAATC 300
QY      1181  CTAAAGTCACTTTTAAAGTGTGAGAAATATGATGATGATGATGATGATGATGATGATG 1240
Db      301  CTAAAGTCACTTTTAAAGTGTGAGAAATATGATGATGATGATGATGATGATGATGATG 360
QY      1241  TCATAGATAAGAACTAAATCAACCTTTTGAAGATCTTATTTGAAGGAGTTGATATATG 1300
Db      361  TCATAGATAAGAACTAAATCAACCTTTTGAAGATCTTATTTGAAGGAGTTGATATATG 420
QY      1301  CCAGAGCTGGATGAGTCTCTGAGGAAATATGTTGGTCCATCTTACTAGATCGCTCCC 1360
Db      421  CCAGAGCTGGATGAGTCTCTGAGGAAATATGTTGGTCCATCTTACTAGATCGCTCCC 480
QY      1361  AGACTCGGCTACAGATAGTGTGATCTCACTCACTCACTCACTCACTCACTCACTCACT 1420
Db      481  AGACTCGGCTACAGATAGTGTGATCTCACTCACTCACTCACTCACTCACTCACTCACT 540
QY      1421  TTATGGAAGGCAGAGACTCACTGAGTCACTGAGTCACTGAGTCACTGAGTCACTGAGT 1480
Db      541  TTATGGAAGGCAGAGACTCACTGAGTCACTGAGTCACTGAGTCACTGAGTCACTGAGT 600
QY      1481  ATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTT 1524
Db      601  ATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTT 644

RESULT 30
LOCUS   CR990494
DEFINITION CR990494 RZPD no. 9016 Homo sapiens CDNA clone RZPDp9016P133 5',
mRNA sequence.
ACCESSION CR990494
VERSION   CR990494.1 GI:68284379
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominiidae; Homo.
          1 (bases 1 to 881)
          Heil, O., Ebert, L., Hennig, S., Henze, S., Radelof, U., Schneider, D.
          and Korn, B.
          Human T-Lymphocytes Library
          Unpublished (2005)
          Contact: Inge Arlart
          RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
          Heubnerweg 6, D-14059 Berlin, Germany
          Email: www.rzpd.de
          RZPD; RZPDp9016P133
          RZPDLIB; (Human T-Lymphocytes) RZPD LIB No. 9016
          http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
          Inge Arlart
          RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
          Heubnerweg 6, D-14059 Berlin, Germany
          Tel: +49 30 32639 100
          Fax: +49 30 32639 111
          www.rzpd.de
          This clone is available from RZPD;
          http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPDp9016P133
          contact RZPD (product support@rzpd.de) for further information.
          Primer name: 663, Primer sequence: CGATAACAATTCACACAG.
          Location/Qualifiers
            1..881
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="RZPDp9016P133"

FEATURES
source

```


Db	781	ATACGGATGGATCC	794	
<div><div>RESULT 31</div><div>BU622228/c</div><div>LOCUS</div><div>DEFINITION</div><div>ACCESSION</div><div>VERSION</div><div>KEYWORDS</div><div>SOURCE</div><div>ORGANISM</div></div>				
<div><div>708 bp</div><div>mRNA</div><div>linear</div><div>EST 23-SEP-2002</div><div>UI-H-FH1-bfu-d-21-0-UI-s1</div><div>NCI CGAP FH1 Homo sapiens cDNA clone</div><div>UI-H-FH1-bfu-d-21-0-UI 3'</div><div>mRNA sequence.</div><div>BU622228</div><div>GI:23288443</div><div>EST.</div><div>Homo sapiens (human)</div><div>Homo sapiens</div><div>Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;</div><div>Hominidae; Homo;</div><div>1 (bases 1 to 708)</div><div>NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.</div><div>National Cancer Institute, Cancer Genome Anatomy Project (CGAP),</div><div>Tumor Gene Index</div><div>Unpublished (1997)</div><div>Contact: Robert Strausberg, Ph.D.</div><div>Email: cgapbs-r@mail.nih.gov</div><div>Tissue Procurement: James Martin</div><div>cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa</div><div>cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa</div><div>DNA Sequencing by: Dr. M. Bento Soares, University of Iowa</div><div>Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu</div><div>Seq primer: M13 FORWARD</div><div>PolyA=Yes.</div></div>				
<div><div>Location/Qualifiers</div><div>1..708</div><div>/organism="Homo sapiens"</div><div>/mol_type="mRNA"</div><div>/db_xref="taxon:9606"</div><div>/clone="UI-H-FH1-bfu-d-21-0-UI"</div><div>/tissue_type="Cell Line"</div><div>/dev_stage="Adult"</div><div>/lab_host="DH10B (Life Technologies)"</div><div>/clone_lib="NCI CGAP FH1"</div><div>/notes="Organ: Chondrosarcoma; Vector: pVT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FH1 is a normalized cDNA library obtained from a cell line derived from grade I chondrosarcoma tissue. The library was constructed and normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pVT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGAATCCGGC. The cell line was provided by Dr. James Martin from the University of Iowa.</div><div>TAG TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1</div><div>Chondrosarcoma</div><div>TAG_LIB=UI-H-FH1</div><div>TAG_SEQ=AGAATCCGGC"</div></div>				
<div><div>ORIGIN</div><div>Query Match 20.5%; Score 639; DB 3; Length 708;</div><div>Best Local Similarity 99.9%; Pred. No. 0;</div><div>Mismatches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</div></div>				
QY	2316	CTTTAAATATATAAATGGGTCAAAATAGAAATGACGATCAGGTGGAGGACTCCAAATATCT	2375	
Db	708	CTTTAAATATATAAATGGGTCAAAATAGAAATGACGATCAGGTGGAGGACTCCAAATATCT	649	
QY	2376	AGCTTCTCGATATCATTTTCTGACTTAGATCTGTGGGCATCCACGGCTGCTCTATGG	2435	
<div><div>FEATURES</div><div>source</div></div>				
<div><div>1..708</div><div>/organism="Homo sapiens"</div><div>/mol_type="mRNA"</div><div>/db_xref="taxon:9606"</div><div>/clone="UI-H-FH1-bfu-d-21-0-UI"</div><div>/tissue_type="Cell Line"</div><div>/dev_stage="Adult"</div><div>/lab_host="DH10B (Life Technologies)"</div><div>/clone_lib="NCI CGAP FH1"</div><div>/notes="Organ: Chondrosarcoma; Vector: pVT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FH1 is a normalized cDNA library obtained from a cell line derived from grade I chondrosarcoma tissue. The library was constructed and normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pVT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGAATCCGGC. The cell line was provided by Dr. James Martin from the University of Iowa.</div><div>TAG TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1</div><div>Chondrosarcoma</div><div>TAG_LIB=UI-H-FH1</div><div>TAG_SEQ=AGAATCCGGC"</div></div>				
<div><div>ORIGIN</div><div>Query Match 20.5%; Score 639; DB 3; Length 708;</div><div>Best Local Similarity 99.9%; Pred. No. 0;</div><div>Mismatches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</div></div>				
QY	2316	CTTTAAATATATAAATGGGTCAAAATAGAAATGACGATCAGGTGGAGGACTCCAAATATCT	2375	
Db	708	CTTTAAATATATAAATGGGTCAAAATAGAAATGACGATCAGGTGGAGGACTCCAAATATCT	649	
QY	2376	AGCTTCTCGATATCATTTTCTGACTTAGATCTGTGGGCATCCACGGCTGCTCTATGG	2435	

```
Db 648 AGCTTCTCGATATGATTTCAATGACTTAGATCGTGTGGGCATCCAGGCTGCTCTATGG 589
QY 2436 AGGATACCTCTCCCTGATGGCAATTAATGACAGAGTTCAGATATCTTCAGGTTGCTATTGC 2495
Db 588 AGGATACCTCTCCCTGATGGCAATTAATGACAGAGTTCAGATATCTTCAGGTTGCTATTGC 529
QY 2496 TGGGGCCCACTGCTCTGAGTCTTCTATGATACAGGATACAGGACCTTATATGGG 2555
Db 528 TGGGGCCCACTGCTCTGAGTCTTCTATGATACAGGATACAGGACCTTATATGGG 469
QY 2556 TCACCTCTGACAGATGAACAGGCTTACTTATGATACAGGATACAGGACCTTATATGGG 2615
Db 468 TCACCTCTGACAGATGAACAGGCTTACTTATGATACAGGATACAGGACCTTATATGGG 409
QY 2616 GTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGTTTCTGATGAGATGTCGA 2675
Db 408 GTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGTTTCTGATGAGATGTCGA 349
QY 2676 TTTTGACATACACAGTATATTAAGTGTGAGGCTGGAAGCCATATGATTT 2735
Db 348 TTTTGACATACACAGTATATTAAGTGTGAGGCTGGAAGCCATATGATTT 289
QY 2736 ACAGATCTATCTCAGGAGACACAGATAGATTCCTGAATCGGAGAACATATGA 2795
Db 288 ACAGATCTATCTCAGGAGACACAGATAGATTCCTGAATCGGAGAACATATGA 229
QY 2796 ACTGCATCTTTGCACTACCTTCAGAAAACCTTGATCAGTATGCTCTAAAGT 2855
Db 228 ACTGCATCTTTGCACTACCTTCAGAAAACCTTGATCAGTATGCTCTAAAGT 169
QY 2856 GATATATTTTGCACCTGTAGACTCTCTGGTATACACTGGCTATTTAAACCAATGAGG 2915
Db 168 GATATATTTTGCACCTGTAGACTCTCTGGTATACACTGGCTATTTAAACCAATGAGG 109
QY 2916 AGGTTTATCAACAGAAACACAGATGATCATCATATTTGATACCTGCCATGTAACA 2975
Db 108 AGGTTTATCAACAGAAACACAGATGATCATCATATTTGATACCTGCCATGTAACA 49
QY 2976 TCTACTCTGAAAATAAATGTGGTGCATG 3005
Db 48 TCTACTCTGAAAATAAATGTGGTGCATG 19

RESULT 32
BUG31054/c
LOCUS
DEFINITION
UI-H-FEI-bdt-j-05-0-UI.s1 NCI CGAP FEI Homo sapiens cDNA clone
BUG31054
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 722)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-x@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library prepared by: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. .722
FEATURES
Source
```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FEI-bdt-j-05-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FEI"
/notes="Organ: Chondrosarcoma; Vector: pYT73-Pac
(Pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI CGAP FEI is a normalized cDNA library
derived from a pool of mRNA obtained from 3 cell lines
from grade II chondrosarcoma tissues. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dr primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor digested with Not I, and cloned directionally
into pYT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
Q18 tail. The sequence tag for this library is
CGCTACGGAC. The cell lines were provided by Dr James
Pacitin from the University of Iowa.
TAG_FISUB=Human grade 2 chondrosarcoma cell line pool
TAG_LIB=UI-H-FEI
TAG_SEQ=CGCTACGGAC"

ORIGIN
Query Match
Best Local Similarity 20.4%; Score 637; DB 3; Length 722;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2484 GGTTCCTATTCCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGA 2543
Db 644 GGTTCCTATTCCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGA 585
QY 2544 AGTTTATATGGGTCACTGACAGGATGAACAGGCTATTAAGTCTGTGGCCAT 2603
Db 584 AGTTTATATGGGTCACTGACAGGATGAACAGGCTATTAAGTCTGTGGCCAT 525
QY 2604 GAAAGCAGAAAAGTTCCCTCTGAAACAAATCGTTTACTGCTCTTACATGTTTCTCGA 2663
Db 524 GAAAGCAGAAAAGTTCCCTCTGAAACAAATCGTTTACTGCTCTTACATGTTTCTCGA 465
QY 2664 TGAGAATGTCCATTTTGGCAGATATATTAAGTCTGTGGATGAGGCTGGAAA 2723
Db 464 TGAGAATGTCCATTTTGGCAGATATATTAAGTCTGTGGATGAGGCTGGAAA 405
QY 2724 GCCATATGATTTACAGATCTATCTCAGGAGACACAGATAGAGTTCCTGAATCGG 2783
Db 404 GCCATATGATTTACAGATCTATCTCAGGAGACACAGATAGAGTTCCTGAATCGG 345
QY 2784 AGAATATGAACTGCATCTTTTGCACATACCAACAACTTGGATCACGATATGC 2843
Db 344 AGAATATGAACTGCATCTTTTGCACATACCAACAACTTGGATCACGATATGC 285
QY 2844 TGCTCTAAAAGTGATATATTTTGGCTGTGTAGACTCTCTGTATATACCTGCTATTT 2903
Db 284 TGCTCTAAAAGTGATATATTTTGGCTGTGTAGACTCTCTGTATATACCTGCTATTT 225
QY 2904 AACCAATGAGGAGTTTAAATCAACAGAAAACAGAAATGATCATCATATTTGATACC 2963
Db 224 AACCAATGAGGAGTTTAAATCAACAGAAAACAGAAATGATCATCATATTTGATACC 165
QY 2964 TGCCATGTAACATCTACTCTGAAAATAAATGTGGTCCCATGAGGGGCTACGGTTTGT 3023
Db 164 TGCCATGTAACATCTACTCTGAAAATAAATGTGGTCCCATGAGGGGCTACGGTTTGT 105
QY 3024 GGTAGTAAATCTAATACCTTAACCCCAATGCTCAAAATGATACATATTTCTTGAGA 3083
Db 104 GGTAGTAAATCTAATACCTTAACCCCAATGCTCAAAATGATACATATTTCTTGAGA 45
```

```
Qy 3084 GACCCAGCAATACCATTAAGATTAATCTAAAAA 3120
Db 44 GACCCAGCAATACCATTAAGATTAATCTAAAAA 8

RESULT 33
BI223892 853 bp mRNA linear EST 11-JUL-2001
LOCUS 602941035F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5104113 5',
DEFINITION mRNA sequence.
ACCESSION BI223892
VERSION BI223892.1 GI:14677336
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 853)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Prepared by: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11251 row: e column: 10
High quality sequence stop: 718.
FEATURES
source
1..853
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5104113"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH MGC 12"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
ORIGIN
Query Match 20.1%; Score 627; DB 2; Length 853;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 677; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1527 TGTGTTTCCCAAGTCAGAGAGAAATTGAGTTATTTTTCCTCTCAAGTCAAAAC 1586
Db 43 TGTGTTTCCCAAGTCAGAGAGAAATTGAGTTATTTTTCCTCTCAAGTCAAAAC 102
Qy 1587 AGGTTTCCGTCATTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAAGCATC 1646
Db 103 AGGTTTCCGTCATTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAAGCATC 162
Qy 1647 CAGTGTGGGTCGCTGCTCCAGTGATTTTCAAGTGCTCTATCAAGAGGAGATAGCAAT 1706
Db 163 CAGTGTGGGTCGCTGCTCCAGTGATTTTCAAGTGCTCTATCAAGAGGAGATAGCAAT 222
Qy 1707 TACCAGTGTGAATGGGAAGTCTTGGCGGCGATCGATCTAATATCCAAGTTGATGAAGT 1766
Db 223 TACCAGTGTGAATGGGAAGTCTTGGCGGCGATCGATCTAATATCCAAGTTGATGAAGT 282
Qy 1767 CAGAGGCTGGTATATTTTGAAGGACCAAGACATCCCTTTAGAGCATCACCTGTACGT 1826
Db 283 CAGAGGCTGGTATATTTTGAAGGACCAAGACATCCCTTTAGAGCATCACCTGTACGT 342
Qy 1827 AGTCAGTTACGTAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACATTC 1886

Db 343 AGTCAGTTACGTAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACATTC 402
Qy 1887 TTGCTGCATCAGTCAGCAGCTGTGACTTCTTTATAAGTAAGTATAGTAACAGAGAAATCC 1946
Db 403 TTGCTGCATCAGTCAGCAGCTGTGACTTCTTTATAAGTAAGTATAGTAACAGAGAAATCC 462
Qy 1947 ACATGTGTGTCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAA 2006
Db 463 ACATGTGTGTCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAA 522
Qy 2007 GGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGCTCTCTTCTCTGACTATCTCTCCAG 2066
Db 523 GGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGCTCTCTTCTCTGACTATCTCTCCAG 582
Qy 2067 AATTTTCTCTTTTGAAGTACTACTGAGATTTACATTTGATGGATGCTCTACAGGCTCA 2126
Db 583 AATTTTCTCTTTTGAAGTACTACTGAGATTTACATTTGATGGATGCTCTACAGGCTCA 642
Qy 2127 TGATCTACAGCCTGGAAAGAAATATCTACTGCTGCTTTCATATATATGATGGTCTCTCAGGT 2186
Db 643 TGATCTACAGCCTGGAAAGAAATATCTACTGCTGCTTTCATATATATGATGGTCTCTCAGGT 702
Qy 2187 GCAGTTGGTGAATAATCG 2204
Db 703 GCAGTTGGTGAATAATCG 720

RESULT 34
CB990233 815 bp mRNA linear EST 01-MAY-2003
LOCUS AGENCOURT_13890379 NIH_MGC_147 Homo sapiens cDNA clone
DEFINITION IMAGE:30347094 5', mRNA sequence.
ACCESSION CB990233
VERSION CB990233.1 GI:30284753
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 815)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM387 row: k column: 07
High quality sequence stop: 587.
FEATURES
source
1..815
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30347094"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 147"
/notes="Organ: placenta; Vector: pBluescriptR; Site 1:
alt-XhoI; Site 2: BamHI; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
```


QY 608 ATTCTCAGAGAAGAACTATTATACAGAAAGAAACGCATTGGACAGTCGGAAATTCGTT 667
 Db 607 ATCTTCGAGAGAGAGACTATTATAGAGAGAGAAACGCATTGGACAGTCGGAAATTCGTT 666
 QY 668 CTTACGATTATCACCAA 684
 Db 667 CTTACGATTATCACCAA 683

RESULT 36
 BG715866
 LOCUS 602675282F1 NIH_MGC_96 695 bp mRNA linear EST 08-MAY-2001
 DEFINITION 602675282F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4797669 5',
 mRNA sequence.
 BG715866
 ACCESSION BG715866.1 GI:13995053
 VERSION
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 695)
 AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10684 row: d column: 22
 High quality sequence stop: 695.
 Location/Qualifiers
 1. .695
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4797669"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 96"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
 (gtcgag); Oligo-dT primed using primer
 5'-TTTCTTTTCTTTTCTTTT-3', size-selected for average
 insert size 2.3 kb and normalized to R0T 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: this is
 a NIH_MGC Library."

ORIGIN
 Query Match 20.0%; Score 623; DB 2; Length 695;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2025 TTGGATTTCAGCAGGTCTCTTCCTGACTATCTCCAGAAATTTCTTTTCAAG 2084
 Db 22 TTGGATTTCAGCAGGTCTCTTCCTGACTATCTCCAGAAATTTCTTTTCAAG 81
 QY 2085 TACTACTGATTATACATGTATGGATGCTCTACAAGCTCATGATCTACAGCTCGAAA 2144
 Db 82 TACTACTGATTATACATGTATGGATGCTCTACAAGCTCATGATCTACAGCTCGAAA 141
 QY 2145 GAAATATCCTACTGTGCTGTTTCATATATGGTGGTCCCTCAGTGCAGTTGGTGAATAATCG 2204

Db 142 GAAATATCCTACTGTGCTGTTTCATATATGGTGGTCTCAGTGCAGTTGGTGAATAATCG 201
 QY 2205 GTTTAAAGGAGTCAAGTATTTCGGCTTGAATACCTCCTAGCTTCTAGGTTATGTGGTTGT 2264
 Db 202 GTTTAAAGGAGTCAAGTATTTCGGCTTGAATACCTCCTAGCTTCTAGGTTATGTGGTTGT 261
 QY 2265 AGTGATAGACAAACAGGGGATCCTGTGCACCGAGGGCTTAAATTTTGAAGGCGCTTTAAATA 2324
 Db 262 AGTGATAGACAAACAGGGGATCCTGTGCACCGAGGGCTTAAATTTTGAAGGCGCTTTAAATA 321
 QY 2325 TAAATGGGTCAAAATGAATTTGACGATCAGTGTGAAGGACTCCAAATATCTAGCTTCTCG 2384
 Db 322 TAAATGGGTCAAAATGAATTTGACGATCAGTGTGAAGGACTCCAAATATCTAGCTTCTCG 381
 QY 2385 ATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCT 2444
 Db 382 ATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCT 441
 QY 2445 CTCCTGATGGCAATTAATGCAGAGTTCAGATATCTTCAGGGTTCCTATTTGCTGGGGCCCC 2504
 Db 442 CTCCTGATGGCAATTAATGCAGAGTTCAGATATCTTCAGGGTTCCTATTTGCTGGGGCCCC 501
 QY 2505 AGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTTATATGGTTCACCCCTGA 2564
 Db 502 AGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTTATATGGTTCACCCCTGA 561
 QY 2565 CAGATATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTC 2624
 Db 562 CTTGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTC 621
 QY 2625 TGAACCAAAATCGTTTACTCTCTTACATGGTTTCTCGATGAGAATGTCCATTTTGCACA 2684
 Db 622 TGAACCAAAATCGTTTACTCTCTTACATGGTTTCTCGATGAGAATGTCCATTTTGCACA 681
 QY 2685 TACCAAGTATATAC 2698
 Db 682 TACCAAGTATATAC 695

RESULT 37
 AI819365/c
 LOCUS AI819365 634 bp mRNA linear EST 07-MAR-2000
 DEFINITION wg61c02.x1 Soares NSF P8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
 IMAGE:2369570 3' similar to TR:075273 O75273 R26984_1 ; mRNA
 sequence.
 ACCESSION AI819365
 VERSION AI819365.1 GI:5438444
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 634)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 699 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 469.
 Location/Qualifiers
 1. .634
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2369570"
 /lab_host="DH10B"

FEATURES
 source

```
/clone_lib="Soares NSF_F8_9W OT_PA_P1.s1"
/note="Organ: pooled; Vector: pMTT30-PacI; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from five
normalized libraries were mixed, and six circles were made
in vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 5 libraries. The pools consisted of the
following libraries and cloneIDs: Soares Nhtsf pool 1:
303384-310919, 323308-325895 Soares N2dHP pool 1:
145032-147335, 147720-148103, 148872-149255 15002 -
150407, 151176-152327 Soares N2tHRG-9W pool 1:
758280-760583, 772104-774407 Soares NHSPA pool 1:
304776-306311, 320133-322823, 326380-326663 Soares
pool 1: 723720-726407, 739060-740999 Subtraction by Bento
Soares and M. Farina 2001
```

ORIGIN

Query Match	Best Local Similarity	19.7%;	Score 614;	DB 1;	Length 634;	Mismatches 614;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2499	GGCCCCAGTCATCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCA	2558							
Db	614	GGCCCCAGTCATCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCA	555							
QY	2559	CCCTGACCAAGAAATGAACAGGGCTATTACTTTAGGATCTGTGCCATGACGAGAAAGTT	2618							
Db	554	CCCTGACCAAGAAATGAACAGGGCTATTACTTTAGGATCTGTGCCATGACGAGAAAGTT	495							
QY	2619	CCCTCTGAAACCAAAATCGTTTACTTGTCTTCTACATGGTTTCTCTGGATGAGAATGTCCATTT	2678							
Db	494	CCCTCTGAAACCAAAATCGTTTACTTGTCTTCTACATGGTTTCTCTGGATGAGAATGTCCATTT	435							
QY	2679	TGCATATACCAAGTATATTACTTGAAGTTTTTGTAGTGAGGCTTGGAAAGCCATATGATTTACA	2738							
Db	434	TGCATATACCAAGTATATTACTTGAAGTTTTTGTAGTGAGGCTTGGAAAGCCATATGATTTACA	375							
QY	2739	GATCTATCTCTCAGAGAGACACAGCATATAGAGTTCTTGAATCGGAGAACCATTTATGAAT	2798							
Db	374	GATCTATCTCTCAGAGAGACACAGCATATAGAGTTCTTGAATCGGAGAACCATTTATGAAT	315							
QY	2799	GCATCTTTTGTGCACTACCTTCAAGAAAACCTTCGGATCAGCATTTGCTGTCTTAAAGTGAT	2858							
Db	314	GCATCTTTTGTGCACTACCTTCAAGAAAACCTTCGGATCAGCATTTGCTGTCTTAAAGTGAT	255							
QY	2859	ATAATTTTGACCTGTGTAGAACCTCTCTGGTATACACTGGCTATTTAACCAATGAGAGG	2918							
Db	254	ATAATTTTGACCTGTGTAGAACCTCTCTGGTATACACTGGCTATTTAACCAATGAGAGG	195							
QY	2919	TTTAAATCAACAGAAAAACACAGAAATGTATCATCAGATTTTGTATACCTGCCATGTAAACATCT	2978							
Db	194	TTTAAATCAACAGAAAAACACAGAAATGTATCATCAGATTTTGTATACCTGCCATGTAAACATCT	135							
QY	2979	ACTCTCTGAAAAATAAATGTGGTGCCATCAGAGGGTCTACGGTTTGTGTAGTAAATCTAAATA	3038							
Db	134	ACTCTCTGAAAAATAAATGTGGTGCCATCAGAGGGTCTACGGTTTGTGTAGTAAATCTAAATA	75							
QY	3039	CCTTTAACCCCAATCCTCTCAAAATCAATATGATACATATTTCTGTAGAGACCCAGCAATACCA	3098							
Db	74	CCTTTAACCCCAATCCTCTCAAAATCAATATGATACATATTTCTGTAGAGACCCAGCAATACCA	15							
QY	3099	TAAGAAATTACTAAA	3112							
Db	14	TAAGAAATTACTAAA	1							

RESULT 38
BQ675006
LOCUS
DEFINITION
BQ675006 957 bp mRNA linear EST 15-JUL-2002
AGSCNCOURT_8303592 NIH_MGC_102 Homo sapiens CDNA clone IMAGE:6274622
5', mRNA sequence.
BQ675006
ACCESSION

```

VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 957)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2456 row, b column: 15
High quality sequence start: 4
High quality sequence stop: 624.
Location/Qualifiers
1..957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6274622"
/tissue_type="epidermoid carcinoma, cell line"
/lab_hosts="DHI0B (phage-resistant)"
/clone_lib="NIH MGC 102"
Note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH MGC Library."

FEATURES
source

```

ORIGIN

Query Match	19.5%; Score 607; DB 3; Length 957;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 607; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1633 AAATATAAAGCATCCAGTGGTGGCGCTGCTCCCAAGTGATTTCAAGTGCCTATCAAA 1692
Db	12 AAATATAAAGCATCCAGTGGTGGCGCTGCTCCCAAGTGATTTCAAGTGCCTATCAAA 71
QY	1693 GAGCAGATAGCAATTACCAGTGGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAAATATC 1752
Db	72 GAGGAGATAGCAATTACCAGTGGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAAATATC 131
QY	1753 CRAGTTGATGAAGTCAGAGCGCTGGTGATATTTTGAAGGCCACCAAGACTCCCCCTTTAGAG 1812
Db	132 CRAGTTGATGAAGTCAGAGCGCTGGTGATATTTTGAAGGCCACCAAGACTCCCCCTTTAGAG 191
QY	1813 CATCATCTGTTACGTAGTCAGTACGTAAATCTCTGGAGAGGTGACAAGSGCTGACTGACCGT 1872
Db	192 CATCACCTGTACGTAGTCAGTACGTAAATCTCTGGAGAGGTGACAAGSGCTGACTGACCGT 251
QY	1873 GGGTACTCACATTTCTTGCTGCATCAGTCAAGCTGTGACTTCTTTTATAGTAAGTATAGT 1932
Db	252 GGGTACTCACATTTCTTGCTGCATCAGTCAAGCTGTGACTTCTTTTATAGTAAGTATAGT 311
QY	1933 AACCAGAAGAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCGAAGATGACCA 1992
Db	312 AACCAGAAGAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCGAAGATGACCA 371
QY	1993 ACTTGGCAAAACAAAGGAAATTTTGGCGCCACCAATTTTGGATTCAGAGGTCCTCTCTCGAC 2052
Db	372 ACTTGGCAAAACAAAGGAAATTTTGGCGCCACCAATTTTGGATTCAGAGGTCCTCTCTCGAC 431

Thu Jun 22 09:04:26 2006

ORIGIN

Query Match 19.2%; Score 598; DB 8; Length 612;
 Best Local Similarity 100.0%; Pred. No. 1.1e-311;
 Matches 598; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 AGGAAATGCAACATGGCAGCAGCAATGGAACAGCAAGCTGGGTGTGGAGATATTTGA 260
 DB 15 AGGAAATGCAACATGGCAGCAGCAATGGAACAGCAAGCTGGGTGTGGAGATATTTGA 74
 QY 261 AACTGCGGAGCTGTGAGGAGAAATTAATCAAGATCGGCTTAATTTGGAGCCCTTTTA 320
 DB 75 AACTGCGGAGCTGTGAGGAGAAATTAATCAAGATCGGCTTAATTTGGAGCCCTTTTA 134
 QY 321 TGTGAGCGGTATTCCTGGAGTCAAGTAAAGCTGTTCGCCGATACAGAGAAATATCA 380
 DB 135 TGTGAGCGGTATTCCTGGAGTCAAGTAAAGCTGTTCGCCGATACAGAGAAATATCA 194
 QY 381 TGGCTACATGATGCTTAAGCAGCACATGATTTTCAATGTTGTGAAGAGAAATGATCCAGA 440
 DB 195 TGGCTACATGATGCTTAAGCAGCACATGATTTTCAATGTTGTGAAGAGAAATGATCCAGA 254
 QY 441 TGGACCTCATTCAGACAGAAATCTATACCTTGCATGCTCTGGTGAGAAACAGAGAAATATCA 500
 DB 255 TGGACCTCATTCAGACAGAAATCTATACCTTGCATGCTCTGGTGAGAAACAGAGAAATATCA 314
 QY 501 ACTGTTTATTTCTGAAATTCCTCAAAATCTATCAATAGACAGCAGCTCTTAATGCTCTCTTG 560
 DB 315 ACTGTTTATTTCTGAAATTCCTCAAAATCTATCAATAGACAGCAGCTCTTAATGCTCTCTTG 374
 QY 561 GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGAATGGAATGCTCTTTCAGGAGAA 620
 DB 375 GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGAATGGAATGCTCTTTCAGGAGAA 434
 QY 621 AGAATCTATTAAGAGAAAGAAACGATGGAACAGCTGGAATGCTCTTTCAGGAGAA 680
 DB 435 AGAATCTATTAAGAGAAAGAAACGATGGAACAGCTGGAATGCTCTTTCAGGAGAA 494
 QY 681 CCAAGGAGTGGAAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAAGTAAAGATGG 740
 DB 495 CCAAGGAGTGGAAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAAGTAAAGATGG 554
 QY 741 AGGCGCCACAAGGATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGT 798
 DB 555 AGGCGCCACAAGGATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGT 612

RESULT 40
 LOCUS CBI53590
 DEFINITION K-EST0211080 B1T694954 Homo sapiens cDNA clone B1T694954-33-D08 5', mRNA linear EST 29-JAN-2003
 mRNA sequence.
 CBI53590
 VERSION CBI53590.1 GI:28138584
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 596)
 Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
 Kim, Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4470
 Email: yongsaung@mail.kribb.re.kr

QY 2053 TATATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTTTACATTTGATGGATG 2112
 DB 432 TATATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTTTACATTTGATGGATG 491
 QY 2113 CTCTACAGGCTCATGATCTACAGCTCGAAAGAAATATCTCTAGTGTGCTTCATATAT 2172
 DB 492 CTCTACAGGCTCATGATCTACAGCTCGAAAGAAATATCTCTAGTGTGCTTCATATAT 551
 QY 2173 GGTGGTCTCAGGTGCGAGTTGGTGAATATCGTTTAAAGAGAGTCAAGATATTCGCGTTG 2232
 DB 552 GGTGGTCTCAGGTGCGAGTTGGTGAATATCGTTTAAAGAGAGTCAAGATATTCGCGTTG 611
 QY 2233 AATACCC 2239
 DB 612 AATACCC 618

RESULT 39
 LOCUS CV024163
 DEFINITION 612 bp mRNA linear EST 20-AUG-2004
 1359 Full length cDNA from the Mammalian Gene Collection Homo
 sapiens cDNA 5' similar to BC030688, mRNA sequence.
 CV024163
 VERSION CV024163.1 GI:51481992
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 612)
 Rual, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S.,
 Driscoll, A., Li, N., Rosenburg, J., Lamesch, P., Vidalain, P.O.,
 Clingingsmith, T.R., Hartley, J.L., Esposito, D., Cheo, D., Moore, T.,
 Simmons, B., Sequerra, R., Bosak, S., Doucette-Stamm, L., Le Peuch, C.,
 Vandenhaute, J., Cusick, M.E., Alcala, J.S., Hill, D.E. and Vidal, M.
 Human ORFeome Version 1.1: a Platform for Reverse Proteomics
 Genome Res. (2004) In press
 Contact: Vidal M
 Marc Vidal Laboratory
 Dana Farber Cancer Institute
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
 Tel: 617 632 5180
 Fax: 617 632 5739
 Email: Marc.Vidal@dfci.harvard.edu
 ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
 results from a PCR reaction using an MGC full-length cDNA as
 template DNA and ORF specific primers
 PCR Primers
 FORWARD: ATGAAATAAATCAGGAAATGCAACATGG
 BACKWARD: TATATCACTTTTATGAGCAGCAATAC
 Insert Length: 612 Std Error: 60.00
 Plate: 11079 row: 01 column: D
 Seq primer: ACTGGCGCTGTTTACAAACGTCGTGACTGGGAAAC
 High quality sequence start: 103
 High quality sequence stop: 611
 POLYA=NO.

FEATURES
 source Location/Qualifiers
 1..612
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="mixed"
 /clone_lib="Full length cDNA from the Mammalian Gene
 Collection"
 /notes="vector: mixed; The ORFs were PCR amplified from the
 MGC (Mammalian Gene Collection) as of April 2004 and
 cloned by recombinational Gateway cloning into pDONR223
 Donor vector. Reference : MGC (Mammalian Gene Collection)
 Program Team, Generation and Initial Analysis of more than
 15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
 2002, 99(26), 16899-16903"

	match	596;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
Qy	54	AGTGC	CGCGTTC	CGCGCT	TGTTGT	CACG	CGCGCG	CGCGCG	CGCGAG	GAAGCACTGCAACC	113
Db	1	AGTGC	CGCGTTC	CGCGCG	TGTTGT	CACG	CGCGCG	CGCGCG	CGCGAG	GAAGCACTGCAACC	60
Qy	114	AGGACC	CGGAGT	GAGCG	CGCGCG	CATGAA	AGCGCG	CAGCGC	CGCTCC	CATAGCGCACGTC	173
Db	61	AGGACC	CGGAGT	GAGCG	CGCGCG	CATGAA	AGCGCG	CAGCGC	CGCTCC	CATAGCGCACGTC	120
Qy	174	GGGAC	CGGTC	CGGCG	CGGGCG	CGGGGG	GAAGAA	TGCACAT	GCGC	GAGCAATGGAAAC	233
Db	121	GGGAC	CGGTC	CGGCG	CGGGCG	CGGGGG	GAAGAA	TGCACAT	GCGC	GAGCAATGGAAAC	180
Qy	234	AGAA	CAGCT	GGTGT	TGAGATA	TTTCAAA	CTCGGACT	GTGAGG	AGATAT	TGAAATCACA	293
Db	181	AGAA	CAGCT	GGTGT	TGAGATA	TTTCAAA	CTCGGACT	GTGAGG	AGATAT	TGAAATCACA	240
Qy	294	GGAT	CGGCT	TAAT	TGGAG	CGCTTT	TATGTT	CAGCGG	GTTT	CCTCGGAGTCAGCTTAAAAA	353
Db	241	GGAT	CGGCT	TAAT	TGGAG	CGCTTT	TATGTT	CAGCGG	GTTT	CCTCGGAGTCAGCTTAAAAA	300

Thu Jun 22 09:04:26 2006

	Matches	587; Conservative	0; Mismatches	0; Indels	0; Gaps	0;
QY	1	AAGTGTCTAAAGCCCTCGAGGCCAAGCCGCTACTCCTCCGCGCTGCTTTTAGTCGG	60			
Db	1	AAGTGTCTAAAGCCCTCGAGGCCAAGCCGCTACTCCTCCGCGCTGCTTTTAGTCGG	60			
QY	61	GTTTGCGCGCTGGTGTGCACCGCGCGCGCGCGCGAGAACCACTCTGAACACGACCG	120			
Db	61	GTTTGCGCGCTGGTGTGCACCGCGCGCGCGCGCGAGAACCACTCTGAACACGACCG	120			
QY	121	GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCGATGCAATGCGAGCAGTGAACAACAG	180			
Db	121	GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCGATGCAATGCGAGCAGTGAACAACAG	180			
QY	181	TCCGGCGCGCGCGCGAGGAAAATGCAATGCGAGCAGTGAACAACAG	240			
Db	181	TCCGGCGCGCGCGCGAGGAAAATGCAATGCGAGCAGTGAACAACAG	240			
QY	241	CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAATATTGAATCACAGGATCG	300			
Db	241	CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAATATTGAATCACAGGATCG	300			
QY	301	CCTAAATTTGGAGCCTTTTTATGTTGAGCGGTATTTCTTGGAGTCACTTAATAAGCTGCTT	360			
Db	301	CCTAAATTTGGAGCCTTTTTATGTTGAGCGGTATTTCTTGGAGTCACTTAATAAGCTGCTT	360			
QY	361	GCCGATACACAGAAAAATATCATGGCTACATGATGGCTAAGGCACCACATGATTTTCATGTTT	420			
- Db	361	GCCGATACACAGAAAAATATCATGGCTACATGATGGCTAAGGCACCACATGATTTTCATGTTT	420			
QY	421	GTGAAGAGGAATGATCCAGATGGACCTCAATCAGACAGAACTTATTACCTGGCCATGCT	480			
Db	421	GTGAAGAGGAATGATCCAGATGGACCTCAATCAGACAGAACTTATTACCTGGCCATGCT	480			
QY	481	GTTGAGAACACAGAAAAATACACTGTTTATTTCTGAAATTTCCCAAACCTATCAATAGAGCA	540			
Db	481	GTTGAGAACACAGAAAAATACACTGTTTATTTCTGAAATTTCCCAAACCTATCAATAGAGCA	540			
QY	541	GCAGTCTTAATGCTCTCTTGGAGCCTCTTTTGGATCTTTTTCAGGC	587			
Db	541	GCAGTCTTAATGCTCTCTTGGAGCCTCTTTTGGATCTTTTTCAGGC	587			
RESULT 43						
LOCUS	DA441156	CTONG2 Homo sapiens cDNA clone CTONG2010006 5', mRNA	586 bp	mRNA	linear	EST 05-NOV-2005
DEFINITION	DA441156	sequence.				
ACCESSION	DA441156.1	GI:80838779				
VERSION	DA441156					
KEYWORDS	EST.	Homo sapiens (human)				
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 586)					
AUTHORS	Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakisaka, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, K., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.					
TITLE	Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes					
JOURNAL	Genome Res. 16 (1), 55-65 (2006)					
PUBMED	16344560					
COMMENT	Contact: Takao Isogai FLJ Project (HRI Team) Helix Research Institute 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: flj-cdna@nifty.com NEIDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB. Location/Qualifiers 1. .587 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="BRTHA2030154" /tissue_type="thalamus" /clone_lib="BRTHA2" /notes="Vector: pME18SFL3"					
FEATURES						
source						
ORIGIN						

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	DATE
Human sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;					
Homnidae; Homo.					
1 (bases 1 to 593)					
Kimura K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa T.,					
Yamashita, K., Yamamoto, J., Sekine, M., Tsutritani, K., Wakaquri, H.,					
Iishi, S., Sugiyama, T., Sato, K., Isono, Y., Irie, R., Kuehida, N.,					
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,					
Murakawa, K., Ishida, S., Iihbashi, T., Takahashi-Fujii, A.					
Tanasei, J., Negai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.					
Diversification of Transcriptional Modulation: Large-scale					
Identification and Characterization of Putative Alternative					
Promoters of Human Genes					
Genome Res. 16 (1), 55-65 (2006)					

```

JOURNAL:
PUBMED:
COMMENT:
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kasarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3985
Email: flj-cdna@nifty.com
NEBO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
FEATURES
Source
1. .593
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="THYMU3036274"
/tissue_type="thymus"
/clone_lib="THYMU3"
/note="Vector: pME18SF1.3"

```

```

ORIGIN
/note="Vector: pME18SFL3"

Query Match      18.8%; Score 586; DB 9; Length 593;
Best Local Similarity 100.0%; Pred. NO. 3.6e-305;
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

248 CTGGGTGTGAGATATTTGAAACTGCGACCTGTGAGGAGAAATTGGAATCACAGATCGG 307

301 CCTAATATCGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTTAAAAAGCTGCTTT 360

308 CCTAATATCGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTTAAAAAGCTGCTTT 367

361 GCCGATACCCAGAAAATATCATGCTACATGATGCTTAAGCCACCATGATTTTCATGTTT 420

368 GCCGATACCCAGAAAATATCATGCTACATGCTTAAGCCACCATGATTTTCATGTTT 427

421 GTGAAGAGGAATGATCCAGATGGACCTTCATTCAGACAGAAATCTATTACCTTGCATGCT 480

TITLE	Murakawa, K., Iehida, S., Iehibaashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S. Differentiation of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL	Genome Res. 16 (1), 55-65 (2006)
PUBMED	16344560
COMMENT	Contact: Takao Isogai FLJ Project (HRI Team) Helix Research Institute 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: flj-cdna@nifty.com NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI) 5; and one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

```

FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end
pass sequencing: RAB.
location/Qualifiers
1..381
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon.9606"
/clone="BRTHA2026924"
/tissue_type="thalamus"
/clone_lib="BRTHA2"
/note="Vector: pME18SFLA"
FEATURES
source

```

[illegible]

241 GGATCGGCTTAATTTGAGCGCTTTTATGTGAGCGGTATTTCTTGGAGTCAGCTTAATAA 300

354 GCTGCTTCGCCATACCAAGAAATATCATGGCTACATGATGGCTAAGCAACCAATGATTT 413

301 GCTGCTTCGCCATACCAAGAAATATCATGGCTACATGATGGCTAAGCAACCAATGATTT 360

414 CATGTTTCTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTATTACCTTGC 473

361 CATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTATTACCTTGC 420

474 CATGCTCGGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAAACTATCAA 533

421 CATGCTCGGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAAACTATCAA 480

534 TAGAGCAGCACTTTAATGCTCTCTTGGAAAGCCCTCTTTTGGATCTTTTTTCAGGCAACACT 593

481 TAGAGCAGCACTTTAATGCTCTCTTGGAAAGCCCTCTTTTGGATCTTTTTTCAGGCAACACT 540

594 GGACTATGGAAATGTTTCTCGAAGAAAGAACTATTAAAGAGA 635

Db	541	GGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAGAGA	582
RESULT 48			
BX645431			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY	730	GTAAAGATGAGGGCCCAAGGATTTAGCGCAACACCTTTAAGGCCCAATCTAGTGGAA	789
Db	1	GTAAAGATGAGGGCCCAAGGATTTAGCGCAACACCTTTAAGGCCCAATCTAGTGGAA	60
QY	790	ACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCGCGCTGATCCAGACTGGATT	849
Db	61	ACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCGCGCTGATCCAGACTGGATT	120
QY	850	GCTTTTATACATCAACGATATTTGGATATCTAACATCGTAACACAGAGAGAAAGGAGA	909
Db	121	GCTTTTATACATCAACGATATTTGGATATCTAACATCGTAACACAGAGAGAAAGGAGA	180
QY	910	CTCATTTATGTCACATAGCTAGCCCAACATGGAAGAGATGCCAGATTCAGCTGGAGTC	969
Db	181	CTCATTTATGTCACATAGCTAGCCCAACATGGAAGAGATGCCAGATTCAGCTGGAGTC	240
QY	970	GCTACCTTTGTTCTCCAAAGAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAA	1029
Db	241	GCTACCTTTGTTCTCCAAAGAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAA	300
QY	1030	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	1089
Db	301	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	360
QY	1090	TCTGAGGTGGAATTTATTTTCATGTTTACATCCCTTATGTTGGAACAAGAGGCGCATTTCA	1149
Db	361	TCTGAGGTGGAATTTATTTTCATGTTTACATCCCTTATGTTGGAACAAGAGGCGCATTTCA	420
QY	1150	TTCCGTTATCTCTAAAAACAGGTACAGCAAAATCTCTAAAGTCACTTTTAAGATGTCGAAATA	1209
Db	421	TTCCGTTATCTCTAAAAACAGGTACAGCAAAATCTCTAAAGTCACTTTTAAGATGTCGAAATA	480
QY	1210	ATGATTGATGCTGAAGGAAGATCATAGATGTCATAGATTAAGAACTAATTAATCAACTTTT	1269
Db	481	ATGATTGATGCTGAAGGAAGATCATAGATGTCATAGATTAAGAACTAATTAATCAACTTTT	540
QY	1270	GAGATTCTTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAAAA	1329
Db	541	GAGATTCTTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAAAA	600
QY	1330	TATGCTTGGTCCATCTCTACTAGATCGCTCCCA	1361
Db	601	TATGCTTGGTCCATCTCTACTAGATCGCTCCCA	632
RESULT 49			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY	730	GTAAAGATGAGGGCCCAAGGATTTAGCGCAACACCTTTAAGGCCCAATCTAGTGGAA	789
Db	1	GTAAAGATGAGGGCCCAAGGATTTAGCGCAACACCTTTAAGGCCCAATCTAGTGGAA	60
QY	790	ACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCGCGCTGATCCAGACTGGATT	849
Db	61	ACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCGCGCTGATCCAGACTGGATT	120
QY	850	GCTTTTATACATCAACGATATTTGGATATCTAACATCGTAACACAGAGAGAAAGGAGA	909
Db	121	GCTTTTATACATCAACGATATTTGGATATCTAACATCGTAACACAGAGAGAAAGGAGA	180
QY	910	CTCATTTATGTCACATAGCTAGCCCAACATGGAAGAGATGCCAGATTCAGCTGGAGTC	969
Db	181	CTCATTTATGTCACATAGCTAGCCCAACATGGAAGAGATGCCAGATTCAGCTGGAGTC	240
QY	970	GCTACCTTTGTTCTCCAAAGAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAA	1029
Db	241	GCTACCTTTGTTCTCCAAAGAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAA	300
QY	1030	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	1089
Db	301	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	360
QY	1090	TCTGAGGTGGAATTTATTTTCATGTTTACATCCCTTATGTTGGAACAAGAGGCGCATTTCA	1149
Db	361	TCTGAGGTGGAATTTATTTTCATGTTTACATCCCTTATGTTGGAACAAGAGGCGCATTTCA	420
QY	1150	TTCCGTTATCTCTAAAAACAGGTACAGCAAAATCTCTAAAGTCACTTTTAAGATGTCGAAATA	1209
Db	421	TTCCGTTATCTCTAAAAACAGGTACAGCAAAATCTCTAAAGTCACTTTTAAGATGTCGAAATA	480
QY	1210	ATGATTGATGCTGAAGGAAGATCATAGATGTCATAGATTAAGAACTAATTAATCAACTTTT	1269
Db	481	ATGATTGATGCTGAAGGAAGATCATAGATGTCATAGATTAAGAACTAATTAATCAACTTTT	540
QY	1270	GAGATTCTTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAAAA	1329
Db	541	GAGATTCTTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAAAA	600
QY	1330	TATGCTTGGTCCATCTCTACTAGATCGCTCCCA	1361
Db	601	TATGCTTGGTCCATCTCTACTAGATCGCTCCCA	632
RESULT 49			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY	730	GTAAAGATGAGGGCCCAAGGATTTAGCGCAACACCTTTAAGGCCCAATCTAGTGGAA	789
Db	1	GTAAAGATGAGGGCCCAAGGATTTAGCGCAACACCTTTAAGGCCCAATCTAGTGGAA	60
QY	790	ACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCGCGCTGATCCAGACTGGATT	849
Db	61	ACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCGCGCTGATCCAGACTGGATT	120
QY	850	GCTTTTATACATCAACGATATTTGGATATCTAACATCGTAACACAGAGAGAAAGGAGA	909
Db	121	GCTTTTATACATCAACGATATTTGGATATCTAACATCGTAACACAGAGAGAAAGGAGA	180
QY	910	CTCATTTATGTCACATAGCTAGCCCAACATGGAAGAGATGCCAGATTCAGCTGGAGTC	969
Db	181	CTCATTTATGTCACATAGCTAGCCCAACATGGAAGAGATGCCAGATTCAGCTGGAGTC	240
QY	970	GCTACCTTTGTTCTCCAAAGAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAA	1029
Db	241	GCTACCTTTGTTCTCCAAAGAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAA	300
QY	1030	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	1089
Db	301	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	360

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 593)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1. .593
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="THYMU3035709"
/tissue_type="thymus"
/clone_lib="THYMU3"
/note="Vector: pME18SFL3"

source

FEATURES

Query Match 18.5%; Score 578; DB 9; Length 593;
Best Local Similarity 100.0%; Pred. No. 7.9e-301;
Matches 578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAGCCTCGAGGCAAGCCGCTGCTACTGCGCGCGCTCTTAGTCCGCGTTCG 67
DB 2 AAGCCTCGAGGCAAGCCGCTGCTACTGCGCGCGCTCTTAGTCCGCGTTCG 61
QY 68 CGCCTGGGTGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTGG 127
DB 62 CGCCTGGGTGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTGG 121
QY 128 GCGCGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 187
DB 122 GCGCGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 181
QY 188 GCGCGCGCGGGAAGAAATGCAATGCGCAGCAGCAATGGAACAGAAACAGTGG 247
DB 182 GCGCGCGCGGGAAGAAATGCAATGCGCAGCAGCAATGGAACAGAAACAGTGG 241
QY 248 TTGAGATATTCGAACTGGGACTGTGAGGAGATATTCGATCAGGATCGGCTAAT 307
DB 242 TTGAGATATTCGAACTGGGACTGTGAGGAGATATTCGATCAGGATCGGCTAAT 301
QY 308 TGAGAGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATA 367
DB 302 TGAGAGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATA 361
QY 368 CCAGAAATATCATGCTACATGATGGCTTAAGGACCAATGATTTTCATGTTGTGAGA 427
DB 362 CCAGAAATATCATGCTACATGATGGCTTAAGGACCAATGATTTTCATGTTGTGAGA 421
QY 428 GGAATGATCCAGATGACCTCAATTCAGACAGAAATCTATTACCTTGGCCATGCTGGGTGAGA 487
DB 422 GGAATGATCCAGATGACCTCAATTCAGACAGAAATCTATTACCTTGGCCATGCTGGGTGAGA 481

QY 488 ACAGAGAAAATACACTGTTTATTTCTGAAATCCCAAACTATCAATAGAGCAGAGTCT 547
DB 482 ACAGAGAAAATACACTGTTTATTTCTGAAATCCCAAACTATCAATAGAGCAGAGTCT 541
QY 548 TAATGCTCTCTTGGAGAGCCTCTTTTGGATCTTTTTCAG 585
DB 542 TAATGCTCTCTTGGAGAGCCTCTTTTGGATCTTTTTCAG 579

RESULT 52
DAL40770
LOCUS DAL40770 BRALZ2 Homo sapiens cDNA clone BRALZ2015423 5', mRNA linear EST 31-OCT-2005
DEFINITION sequence.
ACCESSION DAL40770
VERSION DAL40770.1 GI:78458395
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 629)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1. .629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRALZ2015423"
/tissue_type="alzheimer cortex"
/clone_lib="BRALZ2"
/note="Vector: pME18SFL3"

source

FEATURES

Query Match 18.5%; Score 578; DB 9; Length 629;
Best Local Similarity 99.8%; Pred. No. 7.9e-301;
Matches 628; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 462 CTATTACCTTGGCCATGCTGTTGGTGAGAAACAGAGAAATACACTGTTTATTCGAAATTC 521
DB 1 CTATTACCTTGGCCATGCTGTTGGTGAGAAACAGAGAAATACACTGTTTATTCGAAATTC 60
QY 522 CAAAACTATCAATAGAGCAGCTCTTAATGCTCTTGGAGCCTCTTTGGATCTTTT 581
DB 61 CAAAACTATCAATAGAGCAGCTCTTAATGCTCTTGGAGCCTCTTTGGATCTTTT 120
QY 582 TCAGGCAACACTGGCATGGAATGATTTCTCGAGAGAAAGAACTATTAAAGAGAAAGAAA 641

```
Db 121 TCAGGCAACACTGGACTATGGAATGATTCTCGAGAGAGAACTATTAAAGAGAAAGAAA 180
Qy 642 ACGCATTTGGAACAGTGGAAATGCTTTTACGATTATCACCAGAGAGTGGAACTTCT 701
Db 181 ACGCATTTGGAACAGTGGAAATGCTTTTACGATTATCACCAGAGAGTGGAACTTCT 240
Qy 702 GTTTCAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAGAGATTACGCA 761
Db 241 GTTTCAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAGAGATTACGCA 300
Qy 762 ACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCCAACTACGGATGATCCAAA 821
Db 301 ACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCCAACTACGGATGATCCAAA 360
Qy 822 ATTATGCCCGGTGATCCAGACTGGATTGCTTTTATACATAGCAACGATTTGGATATC 881
Db 361 ATTATGCCCGGTGATCCAGACTGGATTGCTTTTATACATAGCAACGATTTGGATATC 420
Qy 882 TAACATCGTAAACAGAGAGAAAGAGACTCACTATTGTGCAAACTAGCTAGCCAACT 941
Db 421 TAACATCGTAAACAGAGAGAAAGAGACTCACTATTGTGCAAACTAGCTAGCCAACT 480
Qy 942 GGAAGAAGATGCCAGATCAGCTGGAGTGGCTACTTGTTCCTCCAGAGATTTGATAG 1001
Db 481 GGAAGAAGATGCCAGATCAGCTGGAGTGGCTACTTGTTCCTCCAGAGATTTGATAG 540
Qy 1002 ATATTCTGGCTATTGGTGTCCAAAGCTGAAACACTCCCACTGGTGGTAAATTTCT 1061
Db 541 ATATTCTGGCTATTGGTGTCCAAAGCTGAAACACTCCCACTGGTGGTAAATTTCT 600
Qy 1062 TAGAATTTATATGAAGAAATGATGAAT 1090
Db 601 TAGAATTTATATGAAGAAATGATGAAT 629

RESULT 53
DA301389
LOCUS
DEFINITION DA301389 BRHIP2 Homo sapiens cDNA clone BRHIP2016160 5', mRNA linear EST 03-NOV-2005
sequence.
ACCESSION DA301389
VERSION DA301389.1 GI:79169476
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 576)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, P., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yonezawa, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, I., Negai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
1634560
Contract: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
26-7, Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdn@nifty.com
NEO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction;
Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
```

```
FEATURES
source
Location/Qualifiers
1..576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRHIP2016160"
/tissue_type="hippocampus"
/clone_lib="BRHIP2"
/note="Vector: pMB18SFL3"

ORIGIN
Query Match
Best Local Similarity 18.5%; Score 576; DB 9; Length 576;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 14 TCAGAGGCCAAGGGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCGCGCTTGGCGCCCTG 73
Db 1 TCAGAGGCCAAGGGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCGCGCTTGGCGCCCTG 60
Qy 74 GGTTCATACCAGCGCCCGCCGAGGAGCACTGCAACAGGAGCCACTGCAACAGGAGCCGCGCGCC 133
Db 61 GGTTCATACCAGCGCCCGCCGAGGAGCCACTGCAACAGGAGCCGCGCGCCGCGCC 120
Qy 134 GAGCATGAAGCGCGCGCGCGCTCCATAGCGGACGCTCGGGAGCGTCCGGCGCGCGCC 193
Db 121 GAGCATGAAGCGCGCGCGCGCTCCATAGCGGACGCTCGGGAGCGTCCGGCGCGCGCC 180
Qy 194 GGGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACAGAGCAAGAGCTGGTGTGAGA 253
Db 181 GGGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACAGAGCAAGAGCTGGTGTGAGA 240
Qy 254 TATTGAACTCGGAGCTGTGAGGAGAAATATTGAATCAGAGATCGGCTTAATTTGAGC 313
Db 241 TATTGAACTCGGAGCTGTGAGGAGAAATATTGAATCAGAGATCGGCTTAATTTGAGC 300
Qy 314 CTTTTATGTGAGCGGTATTCCTGGAGTACAGCTTAAAGAGCTGTTCCGATACAGAA 373
Db 301 CTTTTATGTGAGCGGTATTCCTGGAGTACAGCTTAAAGAGCTGTTCCGATACAGAA 360
Qy 374 AATATCATGGCTACATGATGGCTAAGGCACACATGATTTTCATGTTTGTGAAGGAATG 433
Db 361 AATATCATGGCTACATGATGGCTAAGGCACACATGATTTTCATGTTTGTGAAGGAATG 420
Qy 434 ATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCATGCTCTGTGTGAGAAAG 493
Db 421 ATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCATGCTCTGTGTGAGAAAG 480
Qy 494 AAAATACACTGTTTATTCTGAAATTCCTCAAACTATCAATAGCAGCAGCTCTTAATGC 553
Db 481 AAAATACACTGTTTATTCTGAAATTCCTCAAACTATCAATAGCAGCAGCTCTTAATGC 540
Qy 554 TCTCTTGAAGCGCTCTTTTGGATCTTTTTCAGGCAA 589
Db 541 TCTCTTGAAGCGCTCTTTTGGATCTTTTTCAGGCAA 576

RESULT 54
DA690898
LOCUS
DEFINITION DA690898 NT2NE2 Homo sapiens cDNA clone NT2NE2010446 5', mRNA linear EST 07-NOV-2005
sequence.
ACCESSION DA690898
VERSION DA690898.1 GI:81154735
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 576)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
```


Thu Jun 22 09:04:26 2006

Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library constructing: HRI,
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1. .576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2NE2010446"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2NE2"
/note="Vector: pME18SFL3; mRNA from NT2 neuron after the
differentiation of NT2 neuronal precursor cells"

ORIGIN

Query Match 18.5%; Score 576; DB 9; Length 576;
Best Local Similarity 100.0%; Pred. No. 9.6e-300;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 222 AGCAATGGAACAGACAGACAGCTGGTCTTGGATATTTGAACCTGGAGGAGAA 281
DB 1 AGCAATGGAACAGACAGACAGCTGGTCTTGGATATTTGAACCTGGAGGAGAA 60
QY 282 TATTGAATTCAGAGATCGGCTAAATTTGGAGCTTTTATGTTGAGCGGTATTCCTGGAG 341
DB 61 TATTGAATTCAGAGATCGGCTAAATTTGGAGCTTTTATGTTGAGCGGTATTCCTGGAG 120
QY 342 TCAGCTTAAAGAGCTGCTGGATACAGAGAAATATCATGCTACATGCTTAAGGC 401
DB 121 TCAGCTTAAAGAGCTGCTGGATACAGAGAAATATCATGCTACATGCTTAAGGC 180
QY 402 ACCACATGATTTTCATGTTGTTGGAAGAGATGATCCAGATGAGACCTCATTCAGACAGAAAT 461
DB 181 ACCACATGATTTTCATGTTGTTGGAAGAGATGATCCAGATGAGACCTCATTCAGACAGAAAT 240
QY 462 CTATTAACCTTGCATGCTGGTGAGACAGAGAAATACACTGTTTATTCGAAATCC 521
DB 241 CTATTAACCTTGCATGCTGGTGAGACAGAGAAATACACTGTTTATTCGAAATCC 300
QY 522 CAAAACTATCAATAGACAGACAGTCTTAATGCTCTTGGAGCTCTTTGGATCTTTT 581
DB 301 CAAAACTATCAATAGACAGACAGTCTTAATGCTCTTGGAGCTCTTTGGATCTTTT 360
QY 582 TCAGGCAACACTGGAATGATGATTTCTCGAGAGAGAACTATTAAGAGAGAGAA 641
DB 361 TCAGGCAACACTGGAATGATGATTTCTCGAGAGAGAACTATTAAGAGAGAGAA 420
QY 642 ACGCATTTGGACACTCGGAATTCCTTACGATATTCACAGAGAGTGAACATTTCT 701
DB 421 ACGCATTTGGACACTCGGAATTCCTTACGATATTCACAGAGAGTGAACATTTCT 480
QY 702 GTTTCACAGCCGGTAGTGGAAATTTATTCAGTAAAGATGGAGGCCACAGAGATTTACGCA 761
DB 481 GTTTCACAGCCGGTAGTGGAAATTTATTCAGTAAAGATGGAGGCCACAGAGATTTACGCA 540

762 ACAACCTTTAAGCCCAATCTAGTGGAACTAGTTG 797
DB 541 ACAACCTTTAAGCCCAATCTAGTGGAACTAGTTG 576

RESULT 55
DA478474 577 bp mRNA linear EST 06-NOV-2005
LOCUS DA478474 DFNES2 Homo sapiens cDNA clone DFNES2008211 5', mRNA
DEFINITION DA478474 sequence.
ACCESSION DA478474
VERSION DA478474.1 GI:80934675
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 577)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library constructing: HRI,
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1. .577
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DFNES2008211"
/cell_type="normal dermal fibroblasts (Neonatal Skin)
(NHDP2564)"
/clone_lib="DFNES2"
/note="Vector: pME18SFL3; primary culture, normal dermal
fibroblasts"

FEATURES

source

ORIGIN

Query Match 18.5%; Score 576; DB 9; Length 577;
Best Local Similarity 100.0%; Pred. No. 9.6e-300; Indels 0; Gaps 0;
Matches 576; Conservative 0; Mismatches 0;

QY 1 AAGTGCTAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCTCTTCTTAGTGCCG 60
DB 2 AAGTGCTAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCTCTTCTTAGTGCCG 61
QY 61 CGTTTCGCGCTCGGTTGTTCACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGCCG 120
DB 62 CGTTTCGCGCTCGGTTGTTCACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGCCG 121
QY 121 GAGTGAGGCGCGCGAGCATGATGAGCGCGCGCGCTTCCATAGCGCACTGCGGACCG 180
DB 122 GAGTGAGGCGCGCGAGCATGATGAGCGCGCGCGCTTCCATAGCGCACTGCGGACCG 181

Identification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
PLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES

Location/Qualifiers
1..575
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRAMY2024215"
/tissue_type="amygdala"
/clone_lib="BRAMY2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 18.4%; Score 575; DB 9; Length 575;
Best Local Similarity 100.0%; Pred. No. 3.3e-299;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 CTTAGTGC CGGCTTCCGCGCTGGGTGTACACGGCGCGCGCGAGGAGCACTGCA 110
DB 1 CTTAGTGC CGGCTTCCGCGCTGGGTGTACACGGCGCGCGCGAGGAGCACTGCA 60

QY 111 ACCAGACCGGAGTGGAGCGCGGCGAGCATGAGCGCGCGCGCGCTCCATGAGCGAC 170
DB 61 ACCAGACCGGAGTGGAGCGCGGCGAGCATGAGCGCGCGCGCGCTCCATGAGCGAC 120

QY 171 GTCCGGACCGTCCGCGCGCGCGCGAGGAGCAATGCAATGCGAGCAGCAATGGA 230
DB 121 GTCCGGACCGTCCGCGCGCGCGCGAGGAGCAATGCAATGCGAGCAGCAATGGA 180

QY 231 AACAGAACGCTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATC 290
DB 181 AACAGAACGCTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATC 240

QY 291 ACAGGATCGGCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAA 350
DB 241 ACAGGATCGGCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAA 300

QY 351 AAAGTGTCTTCCGATACAGAGAAATATCATGGCTACATGAGCTTAAGCACCACATGA 410
DB 301 AAAGTGTCTTCCGATACAGAGAAATATCATGGCTACATGAGCTTAAGCACCACATGA 360

QY 411 TTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTACAGACAGATCTATTACCT 470
DB 361 TTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTACAGACAGATCTATTACCT 420

QY 471 TCCCATGTCTGGTGAGAACAGAGAAATACACTGCTTTTNTTCTGAAATTTCCCAAACTAT 530
DB 421 TGCCATGTCTGGTGAGAACAGAGAAATACACTGCTTTTNTTCTGAAATTTCCCAAACTAT 480

QY 531 CAATGAGCAGCAGTCTTAATGCTCTCTTTGAGCGCTCTTTTGATCTTTTTCAGGCAAC 590
DB 481 CAATGAGCAGCAGTCTTAATGCTCTCTTTGAGCGCTCTTTTGATCTTTTTCAGGCAAC 540

QY 591 ACTGGACTATGAAATGTTATTTTCGAGAGAGAAAC 625
DB 541 ACTGGACTATGAAATGTTATTTTCGAGAGAGAAAC 575

RESULT 58

DA769628 OCBF2 Homo sapiens cDNA clone OCBF2002128 5', mRNA
573 bp
sequence.
DA769628 OCBF2 Homo sapiens (human)
DA769628 1 GI:82085548
EST.
KEYWORDS
SOURCE
ORGANISM

REFERENCE

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.

TITLE

Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)

JOURNAL

PUBMED
16344560
Contact: Takao Isogai
PLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com

COMMENT

NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES

Location/Qualifiers
1..573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OCBF2002128"
/tissue_type="brain"
/dev_stage="fetal"
/clone_lib="OCBF2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 18.4%; Score 573; DB 9; Length 573;
Best Local Similarity 100.0%; Pred. No. 4.1e-298;
Matches 573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 CTTAGTGC CGGCTTCCGCGCTGGGTGTACACGGCGCGCGCGAGGAGCACTGCA 110
DB 1 CTTAGTGC CGGCTTCCGCGCTGGGTGTACACGGCGCGCGCGAGGAGCACTGCA 60

QY 111 ACCAGACCGGAGTGGAGCGCGGCGAGCATGAGCGCGCGCGCTCCATGAGCGAC 170
DB 61 ACCAGACCGGAGTGGAGCGGCGAGCATGAGCGCGCGCGCTCCATGAGCGAC 120

QY 171 GTCCGGACCGTCCGCGCGCGCGCGAGGAGCAATGCAATGCGAGCAGCAATGGA 230
DB 121 GTCCGGACCGTCCGCGCGCGCGCGAGGAGCAATGCAATGCGAGCAGCAATGGA 180

QY 231 AACAGAACGCTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATC 290
DB 181 AACAGAACGCTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATC 240

QY 291 ACAGGATCGGCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAA 350
DB 241 ACAGGATCGGCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAA 300

```
QY 351 AAAGCTGCTTGGCCGATACAGAAATATCATGGCTACATGCTAAGGCACCATGA 410
Db 301 AAAGCTGCTTGGCCGATACAGAAATATCATGGCTACATGCTAAGGCACCATGA 360
QY 411 TTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCT 470
Db 361 TTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCT 420
QY 471 TGCATGCTGCTGAGAACAGAGAAATACACTGTTTATCTGAAATTCACCAAACTAT 530
Db 421 TGCATGCTGCTGAGAACAGAGAAATACACTGTTTATCTGAAATTCACCAAACTAT 480
QY 531 CAATAGAGCAGCAGTCTTAATGCTCTCTGGAGGCCCTTTTGGATCTTTTTCAGGCAAC 590
Db 481 CAATAGAGCAGCAGTCTTAATGCTCTCTGGAGGCCCTTTTGGATCTTTTTCAGGCAAC 540
QY 591 ACTGACTATGGAATGATTTCTCGAGAGGAAGA 623
Db 541 ACTGACTATGGAATGATTTCTCGAGAGGAAGA 573

RESULT 59
DA931907
LOCUS
DEFINITION
DA931907 SPLEN2 Homo sapiens cDNA clone SPLN200115 5', mRNA
ACCESSION
DA931907
VERSION
DA931907.1 GI:82063318
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 574)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdn@nifty.com
NEPO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction.
Helix Research Institute (HRI); 5'-end one pass sequencing. HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation, 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1..574
/organism="Homo sapiens"
/mol_type="mRNA"
/nci_xref="taxon:9606"
/clone="SPLN200115"
/tissue_type="spleen"
/clone_lib="SPLN2"
/note="Vector: pME18SPL3"

FEATURES
source
Query Match 18.4%; Score 573; DB 9; Length 574;
Best Local Similarity 100.0%; Pred. No. 4.1e-298;

ORIGIN
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Tel: 81-438-52-3986
```

```
Matches 573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 AAAGCTTCCGAGGCCAAGCCGCTGCTACTGCCGCCGCTGCTTCTTAGTGCAGCTCCG 67
Db 2 AAAGCTTCCGAGGCCAAGCCGCTGCTACTGCCGCCGCTGCTTCTTAGTGCAGCTCCG 61
QY 68 CCGCTGGGTGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 61
Db 62 CCGCTGGGTGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 127
QY 128 GCGCGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 121
Db 122 GCGCGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 187
QY 188 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 181
Db 182 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 247
QY 248 TTGAGATATTGAAACTGCGGACTGTGAGGAGCAATATTGAATCAGAGTCGGCTAAAT 241
Db 242 TTGAGATATTGAAACTGCGGACTGTGAGGAGCAATATTGAATCAGAGTCGGCTAAAT 307
QY 308 TGGAGCCTTTTATGTTGAGCGGTATCTCTGGAGTCAGCTTTAAAGCTGTTGCCGATA 301
Db 302 TGGAGCCTTTTATGTTGAGCGGTATCTCTGGAGTCAGCTTTAAAGCTGTTGCCGATA 367
QY 368 CCAGAAATATCATGCTGCTACATGATGGCTAAGGCCACCATGATTTTGTGGAAGA 427
Db 362 CCAGAAATATCATGCTGCTACATGATGGCTAAGGCCACCATGATTTTGTGGAAGA 421
QY 428 GGAATGATCCAGATGACCTCTATTCCAGACAGAACTATTACCTGCGATGCTGGTGA 487
Db 422 GGAATGATCCAGATGACCTCTATTCCAGACAGAACTATTACCTGCGATGCTGGTGA 481
QY 488 ACAGAGAAATACACTGTTTATCTGAAATTCCTGAGTCAGCTTTAAAGCTGTTGCCG 547
Db 482 ACAGAGAAATACACTGTTTATCTGAAATTCCTGAGTCAGCTTTAAAGCTGTTGCCG 541
QY 548 TAATGCTCTCTTGGAGGCTCTTTTGGATCTTT 580
Db 542 TAATGCTCTCTTGGAGGCTCTTTTGGATCTTT 574

RESULT 60
DA931907
LOCUS
DEFINITION
DA931907 PLACE7 Homo sapiens cDNA clone PLACE7010347 5', mRNA
ACCESSION
DA931907
VERSION
DA931907.1 GI:82360681
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 573)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Tel: 81-438-52-3986
```



```

/tissue_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clonelib="GRN ES"
/notes="Oligo dT primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p36) maintained in feeder-free conditions"

```

ORIGIN	Query Match	Best Local Similarity	Matches 561;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
		18.0%;	Score 561;	DB 8;	Length 678;						
QY	25	GGCGCGCTGCTACTGGCGCGCTGCTCTTAGTGGCGCGTTCGGCGCTGGGTTGTCAACG	84								
Db	1	GGCGCGCTGCTACTGGCGCGCGCTGCTCTTAGTGGCGCGTTCGGCGCTGGGTTGTCAACG	60								
QY	85	GGCGCGCGCGCGGAGAGCCACTGCAACAGAGACCGAGTGGAGCGGCGCGCAGCATGAAG	144								
Db	61	GGCGCGCGCGCGGAGAGCCACTGCAACAGAGACCGAGTGGAGCGGCGCGCAGCATGAAG	120								
QY	145	GGCGCGAGCGCGCTCCATAGCGACGTCGCGGACGTCGCGGCGGCGCGCGGGGGAAGGA	204								
Db	121	GGCGCGAGCGCGCTCCATAGCGACGTCGCGGACGTCGCGGCGGCGCGCGGGGGAAGGA	180								
QY	205	AAATGCAACATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAAACT	264								
Db	181	AAATGCAACATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAAACT	240								
QY	265	GGGACGTGCGAGAGAAATATGCAATCAGCAGATCGGCCTAAATTTGGAGCCTTTTATGTT	324								
Db	241	GGGACGTGCGAGAGAAATATGCAATCAGCAGATCGGCCTAAATTTGGAGCCTTTTATGTT	300								
QY	325	GAGCGGTAATTCCTGGAGTCAGCTAAAGAGCTGCTTCGCATACCAAGAAATATCATGCG	384								
Db	301	GAGCGGTAATTCCTGGAGTCAGCTAAAGAGCTGCTTCGCATACCAAGAAATATCATGCG	360								
QY	385	TACATGATGGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGATATCCAGATGGA	444								
Db	361	TACATGATGGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGATATCCAGATGGA	420								
QY	445	CCTCATTCGACAGAAATCTATTCCTTGGCATGCTCTGCTGAGAACAGAGAAATATCACTG	504								
Db	421	CCTCATTCGACAGAAATCTATTCCTTGGCATGCTCTGCTGAGAACAGAGAAATATCACTG	480								
QY	505	TTTTATCTCGAAATTCCTTGGCATGCTCAATAGAGCAGCAGCTTAACTCTCTTTGGAAG	564								
Db	481	TTTTATCTCGAAATTCCTTGGCATGCTCAATAGAGCAGCAGCTTAACTCTCTTTGGAAG	540								
QY	565	CCTCTTTTGGATCTTTTTCAG	585								
Db	541	CCTCTTTTGGATCTTTTTCAG	561								

RESULT 64	AW303607/c	linear	673 bp	mRNA	EST 18-JAN-2000
LOCUS	AW303607				
DEFINITION	X218121.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone IMAGE:2813759 3' similar to TR:075273 R36984_1 ; mRNA sequence.				
ACCESSION	AW303607				
VERSION	AW303607.1		GI:6713296		
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 673)				
AUTHORS	NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap .				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D.				

Email: csapps@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 465.
Location/Qualifiers
1. .673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2813759"
/lab_host="DH10B"
/clone_lib="Soares NFL T CBC S1"
/note="Organ: pooled; Vector: p773D-Pact; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (fetal lung NBHL19W, testis NHT, and
B-cell NCI CGAP GCBI) were mixed, and ss circles were made
in vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

Query Match	17.9%	Score 558	DB 7	Length 673
Best Local Similarity	99.7%	Pred. No. 5.6e-290		
Matches 658	Conservative	0	Mismatches 2	Indels 0
Gaps				
QY	2450	TGATGGCAATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGGCCCAAGTCA	2509	
DB	660	TGATGGCAATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCACCAAGTCA	601	
QY	2510	CTCTGTGGATCTCTTATGATACAGGATACAGGAACGTTATATGGGTCAACCTGACCGA	2569	
DB	600	CTCTGTGGATCTCTTATGATACAGGATACAGGAACGTTATATGGGTCAACCTGACCGA	541	
QY	2570	ATGAACAGGGCTATTACTTTAGGATCTGTGGCCATCGAAGCAAAAGTTGCCCTCTGAAC	2629	
DB	540	ATGAACAGGGCTATTACTTTAGGATCTGTGGCCATCGAAGCAAAAGTTGCCCTCTGAAC	481	
QY	2630	CAAAATCGTTTACTGCTCTTACATGGTTTCCTGGATGAGAATGTCTATTTGCAATACCA	2689	
DB	480	CAAAATCGTTTACTGCTCTTACATGGTTTCCTGGATGAGAATGTCTATTTGCAATACCA	421	
QY	2690	GTATATTACTGAGTGTGTTTTTGTAGCGGGCTGGAAAGCCATGATGATTTACAGATCTATCCTC	2749	
DB	420	GTATATTACTGAGTGTGTTTTTGTAGCGGGCTGGAAAGCCATGATGATTTACAGATCTATCCTC	361	
QY	2750	AGGAGAGACACAGATAAAGATGTTCTGAAATCGGGGAAACATTTATGAACTGCACTTTTTCG	2809	
DB	360	AGGAGAGACACAGCATAAAGATGTTCTGAAATCGGGGAAACATTTATGAACTGCACTTTTTCG	301	
QY	2810	ACTACCTTCAAGAAACCTCTGGATCAAGTATGCTGCTCTAAAGTGATATAATTTTGAC	2869	
DB	300	ACTACCTTCAAGAAACCTCTGGATCAAGTATGCTGCTCTAAAGTGATATAATTTTGAC	241	
QY	2870	CTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAATCAACA	2929	
DB	240	CTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAATCAACA	181	
QY	2930	GAAACACAGAAATTTGATCATCATTTTTGATACCTGCCATGATACACTCTCTCTGAAA	2989	
DB	180	GAAACACAGAAATTTGATCATCATTTTTGATACCTGCCATGATACACTCTCTCTGAAA	121	
QY	2990	TAAATGTGGTGCCATCGAGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAAACCCA	3049	
DB	120	TAAATGTGGTGCCATCGAGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAAACCCA	61	
QY	3050	CATGCTCAAAATCAAAATGATACATATTTCTGTGAGAGCCAGCAATACCATTAAGATTACT	3109	
DB	60	CATGCTCAAAATCAAAATGATACATATTTCTGTGAGAGCCAGCAATACCATTAAGATTACT	1	

Page 48

ORIGIN

Query Match 17.8%; Score 556; DB 9; Length 558;
 Best Local Similarity 100.0%; Pred. No. 6.7e-289;
 Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGTGTAAAGGCTCGAGGCGAAGCGCGCTGCTACTGCGCGCGCTGCTTCTAGTGGCG 60
 DB 3 AAGTGTAAAGGCTCGAGGCGAAGCGCGCTGCTACTGCGCGCGCTGCTTCTAGTGGCG 62

QY 61 CGTTCCGCCCTCGGCTGTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 DB 63 CGTTCCGCCCTCGGCTGTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122

QY 121 GAGTGGAGCG 180
 DB 123 GAGTGGAGCG 182

QY 181 TCCGGCG 240
 DB 183 TCCGGCG 242

QY 241 CTGGGCTGTGAGATATTTGAAATCTCGGACTCTGAGGAGATATTTGAAATCTGAGGATCGG 300
 DB 243 CTGGGCTGTGAGATATTTGAAATCTCGGACTCTGAGGAGATATTTGAAATCTGAGGATCGG 302

QY 301 CTTAAATTCGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360
 DB 303 CTTAAATTCGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 362

QY 361 GCCGATACAGAGAAATATCATGGCTACATGATGGCTAAGGACACATGATTTTCATGTTT 420
 DB 363 GCCGATACAGAGAAATATCATGGCTACATGATGGCTAAGGACACATGATTTTCATGTTT 422

QY 421 GTGAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTCTATTACCTTGGCATGTCT 480
 DB 423 GTGAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTCTATTACCTTGGCATGTCT 482

QY 481 GGTGAGAACAGAGAGAAATACACTGTTTTATTCGAAATTCGAAATTCGAAATTCGAAATTCG 540
 DB 483 GGTGAGAACAGAGAGAAATACACTGTTTTATTCGAAATTCGAAATTCGAAATTCGAAATTCG 542

QY 541 GCAGTCTTAATGCTCT 556
 DB 543 GCAGTCTTAATGCTCT 558

RESULT 67
 BG163397
 LOCUS
 DEFINITION 602338360F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4446375 5', mRNA sequence.

ACCESSION BG163397
 VERSION BG163397.1
 KEYWORDS GI:12670100
 SOURCE EST.
 ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 788)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: the I.M.A.G.E. Consortium (LNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov

Plate: LLAM10225 row: g column: 16
 High quality sequence start: 6
 High quality sequence stop: 742.
 Location/Qualifiers
 1. .788
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4446375"
 /tissue_type="hypernephroma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 89"
 /notes="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 17.8%; Score 555; DB 2; Length 788;
 Best Local Similarity 99.8%; Pred. No. 2.4e-288;
 Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2049 TGACTATACCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGG 2108
 DB 4 TGACTATACCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGG 63

QY 2109 GATGCTCTACAAGCCTCATGATCTACAGCCTCGAAGAAATATCTACTGTGCTGTTTCAT 2168
 DB 64 GATGCTCTACAAGCCTCATGATCTACAGCCTCGAAGAAATATCTACTGTGCTGTTTCAT 123

QY 2169 ATATGTGTGCTTCAGGTGAGTGTGTAATATCGTTTAAAGAGTCAAGTATTTCCG 2228
 DB 124 ATATGTGTGCTTCAGGTGAGTGTGTAATATCGTTTAAAGAGTCAAGTATTTCCG 183

QY 2229 CTTGAATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCCTG 2288
 DB 184 CTTGAATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCCTG 243

QY 2289 TCACCGAGGCTTAAATTTGAGGGCGCTTTAAATATATAAATGGGTCAATAGAAATTTGA 2348
 DB 244 TCACCGAGGCTTAAATTTGAGGGCGCTTTAAATATATAAATGGGTCAATAGAAATTTGA 303

QY 2349 CGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCG 2408
 DB 304 CGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCG 363

QY 2409 TGTGGGCATCCAGCGTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAG 2468
 DB 364 TGTGGGCATCCAGCGTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAG 423

QY 2469 GTCAAGATATCTTCAGGGTGTGCTATTGCTGGGGCCCCCAGTCACTCTGTGGATCTTCTATGA 2528
 DB 424 GTCAAGATATCTTCAGGGTGTGCTATTGCTGGGGCCCCCAGTCACTCTGTGGATCTTCTATGA 483

QY 2529 TACAGGATACACGGAACGTTTATATGAGGTCACTGATGACGAAATGAACAGGGCTATTACTT 2588
 DB 484 TACAGGATACACGGAACGTTTATATGAGGTCACTGATGACGAAATGAACAGGGCTATTACTT 543

QY 2589 AGGATCTGTGGCCATCAAGCAGAAAAGTTCCCTCTCTGAACCAAAATCGTTTACTGCTCTT 2648
 DB 544 AGGATCTGTGGCCATCAAGCAGAAAAGTTCCCTCTCTGAACCAAAATCGTTTACTGCTCTT 603

QY 2649 ACATGG 2654
 DB 604 ACATGG 609

RESULT 68
 DB074943
 LOCUS
 DEFINITION DB074943 TEST14 Homo sapiens cDNA clone TEST14018676 5', mRNA sequence.

DB074943
DB074943.1 GI:83100418
EST.
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 553)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1. 553
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TEST14018676"
/tissue="testis"
/clone_lib="TEST14"
/note="Vector: pME18SFL3"

Query Match
Best Local Similarity 17.7%; Score 553; DB 9; Length 553;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 AGTGGCGGTCCTCCCGCTGGGTGTACCGCGCGCGCGCGAGGAGCCACTGCAACC 113
DB 1 AGTGGCGGTCCTCCCGCTGGGTGTACCGCGCGCGCGCGAGGAGCCACTGCAACC 60

QY 114 AGACGCGGATGAGCGCGCGCAGCATGAAGCGCGCGAGCGCCCTCATAGCGCAGTC 173
DB 61 AGACGCGGATGAGCGCGCGCAGCATGAAGCGCGAGCGCGCGCCCTCATAGCGCAGTC 120

QY 174 GGGACGCTCGGCGCGCGCGCGGCGGGAAGGAAATGCAATGCGCAGCATGGAAC 233
DB 121 GGGACGCTCGGCGCGCGCGCGGCGGGAAGGAAATGCAATGCGCAGCATGGAAC 180

QY 234 AGAAGCTGGGTGTGAGATTTGAACTGCGGACTGTGAGAGAGATTTGAATCA 293
DB 181 AGAAGCTGGGTGTGAGATTTGAACTGCGGACTGTGAGAGAGATTTGAATCA 240

QY 294 GGAATGGCTAAATPAGGACCTTTTATGTGAGCGGTATTCCTGAGTCACTTAA 353
DB 241 GGAATGGCTAAATPAGGACCTTTTATGTGAGCGGTATTCCTGAGTCACTTAA 300

QY 354 GCTGCTGCCGATACCAAGAAATATCATGCTCATGATGGCTTAAGGACCATGATTT 413
DB 301 GCTGCTGCCGATACCAAGAAATATCATGCTCATGATGGCTTAAGGACCATGATTT 360

QY 414 CATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTTACCTTGC 473

DB 361 CATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTTACCTTGC 420

QY 474 CATGCTGTGTGAGACAGAGAAAATACACTGTTTTATTCTGAAATCCCAAACTCA 533
DB 421 CATGCTGTGTGAGACAGAGAAAATACACTGTTTTATTCTGAAATCCCAAACTCA 480

QY 534 TAGACGACGAGCTTTAATGCTCTCTTGGAGACCTCTTTTGATCTTTTCAGGCACT 593
DB 481 TAGACGACGAGCTTTAATGCTCTCTTGGAGACCTCTTTTGATCTTTTCAGGCACT 540

QY 594 GCATCATGGAATG 606
DB 541 GCATCATGGAATG 553

RESULT 69
LOCUS DA268082
DEFINITION DA268082 BRCAN2 Homo sapiens cDNA clone BRCAN2021908 5', mRNA
sequence.
ACCESSION DA268082
VERSION DA268082.1 GI:78258983
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 566)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1. 553
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TEST14018676"
/tissue="testis"
/clone_lib="TEST14"
/note="Vector: pME18SFL3"

Query Match
Best Local Similarity 17.7%; Score 552; DB 9; Length 566;
Matches 552; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 CCAAGGCGCTGCTACTGCGCGCGCTGCTTTAGTCCGCGTTCGCGCTGGTTGTC 80
DB 1 CCAAGGCGCTGCTACTGCGCGCGCTGCTTTAGTCCGCGCTGGTTGTC 60

```
QY 81 ACCGGCGCGCGCGGAGGAGCCACTGCAACAGGACCGGAGTGGAGGCGGCGGAGCAT 140
Db 61 ACCGGCGCGCGCGCGGAGGAGCCACTGCAACAGGACCGGAGTGGAGGCGGCGGAGCAT 120
QY 141 GAAGCGCGCGCGCGCGGCTCCATAGCGCAGCTCGGGACGGTCCGGGCGGGCGCGGGGGA 200
Db 121 GAAGCGCGCGCGCGCGGCTCCATAGCGCAGCTCGGGACGGTCCGGGCGGGCGCGGGGGA 180
QY 201 AGGAAATGCAATGCGAGCAGCAAGATGGAAACAGACAGCTGGGTGTTGAGATATTGA 260
Db 181 AGGAAATGCAATGCGAGCAGCAAGATGGAAACAGACAGCTGGGTGTTGAGATATTGA 240
QY 261 AACTGCGGACTGTGAGGAGAAATATTGAATCAAGGATCGGCCTAAATTTGGAGCCTTTTGA 320
Db 241 AACTGCGGACTGTGAGGAGAAATATTGAATCAAGGATCGGCCTAAATTTGGAGCCTTTTGA 300
QY 321 TGTGAGCGGTATTCCTGAGTCAGCTTAAAGCTGCTTGGCGATACCCAGAAATATCA 380
Db 301 TGTGAGCGGTATTCCTGAGTCAGCTTAAAGCTGCTTGGCGATACCCAGAAATATCA 360
QY 381 TGGCTACATGATGGCTAAGGACCAACATGATTTTCATGTTGTGAAGAGGAATGATCCAGA 440
Db 361 TGGCTACATGATGGCTAAGGACCAACATGATTTTCATGTTGTGAAGAGGAATGATCCAGA 420
QY 441 TGGACCTCATTCAGACAGAAATCTATTACCTTGCCATGCTGTTGGTGAGAACAGAGAAATAC 500
Db 421 TGGACCTCATTCAGACAGAAATCTATTACCTTGCCATGCTGTTGGTGAGAACAGAGAAATAC 480
QY 501 ACTGTTTTATCTGAAATTTCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTG 560
Db 481 ACTGTTTTATCTGAAATTTCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTG 540
QY 561 GAAGCCTCTTTT 572
Db 541 GAAGCCTCTTTT 552
```

```
RESULT 70
CN427230
LOCUS CN427230 553 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000532204093 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN427230
VERSION CN427230.1 GI:47414824
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 553)
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Brandenberger R
Regenerative Medicine
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 553 Std Error: 0.00.
Location/Qualifiers
FEATURES
source
1..553
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone_lib="GRN_ES"
```

```
/note="oligo dT primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"

ORIGIN
Query Match 17.6%; Score 548; DB 8; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.5e-284;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AGCGCGCTGCTACTGCGCGCGCTTCTTAGTGCAGCTTCGCGCGCTTGTCTCACC 83
Db 6 AGCGCGCTGCTACTGCGCGCGCTTCTTAGTGCAGCTTCGCGCGCTTGTCTCACC 65
QY 84 GCGCGCGCGCGCGGAGGACCACTGCAACAGGACCGGAGTGGAGGCGGCGGAGGAA 143
Db 66 GCGCGCGCGCGCGGAGGACCACTGCAACAGGACCGGAGTGGAGGCGGCGGAGGAA 125
QY 144 GCGCGGAGGCGCGCTTCATAGCGCAGCTCGGGACGGTCCGGGCGGGCGGGGGAAGG 203
Db 126 GCGCGGAGGCGCGCTTCATAGCGCAGCTCGGGACGGTCCGGGCGGGCGGGGGAAGG 185
QY 204 AAAATGCAATGCGGAGCAATGAAACAGACAGCTGGGTGTTGAGATATTGAAAC 263
Db 186 AAAATGCAATGCGGAGCAATGAAACAGACAGCTGGGTGTTGAGATATTGAAAC 245
QY 264 TCGGACTGTGAGGAGAAATTTGAATCAAGGATCGGCCTAAATTTGGAGCCTTTTATGT 323
Db 246 TCGGACTGTGAGGAGAAATTTGAATCAAGGATCGGCCTAAATTTGGAGCCTTTTATGT 305
QY 324 TGAGCGGTATTCCTGAGTCAGCTTAAAGCTGCTTCCGATACCCAGAAATATCATGG 383
Db 306 TGAGCGGTATTCCTGAGTCAGCTTAAAGCTGCTTCCGATACCCAGAAATATCATGG 365
QY 384 CTACATGATGGCTAAGGACCAACATGATTTTCATGTTGTGAAGAGGAATGATCAGATGG 443
Db 366 CTACATGATGGCTAAGGACCAACATGATTTTCATGTTGTGAAGAGGAATGATCAGATGG 425
QY 444 ACCTCATTCAGACAGAAATCTATTACCTTGCCATGCTGTCGAGAACAGAGAAATACACT 503
Db 426 ACCTCATTCAGACAGAAATCTATTACCTTGCCATGCTGTCGAGAACAGAGAAATACACT 485
QY 504 GTTTTATTCGAAATTTCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGAA 563
Db 486 GTTTTATTCGAAATTTCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGAA 545
QY 564 GCCTCTTT 571
Db 546 GCCTCTTT 553

RESULT 71
CN427227
LOCUS CN427227 612 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000532648202 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN427227
VERSION CN427227.1 GI:47414821
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 612)
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
```

```

230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@eron.com
Insert Length: 612 Std Error: 0.00.
Location/Qualifiers
1. .612
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon.9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone_lib="GRN ES"
/note="oligo dr primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions".

```

Query Match	17.6%;	Score 548;	DB 8;	Length 612;
Best Local Similarity	99.8%;	Pred. No. 1.5e-284;		
Matches 539;	Conservative	0;	Mismatches 1;	Indels 0;
QY	24	AGGCCGCTGCTACTGCCCGCGCTGCTTCTTAGTCCCGGTTCCGCCGCTGGGTGTCACC	83	
Db	14	AGGCCGCTGCTACTGCCCGCGCTGCTTCTTAGTCCCGGTTCCGCCGCTGGGTGTCACC	73	
QY	84	GGCGCGCGCGCGAGGAAGCCACTGCAACAGAGNCCGGAGTGGAGCGCGCGCATGAATGAA	143	
Db	74	GGCGCGCGCGCGAGGAAGCCACTGCAACAGAGNCCGGAGTGGAGCGCGCGCATGAATGAA	133	
QY	144	GGCGCGCAGGCGCGCTCCATAGCCAGCTCGGAGNCGGTCGGCGCGGGCGGGGGGAAGG	203	
Db	134	GGCGCGCAGGCGCGCTCCATAGCCAGCTCGGAGNCGGTCGGCGCGGGCGGGGGGAAGG	193	
QY	204	AAATGCAACATGCGACGAGCAATGGAAAACAGAACAGCTGGGTGTTGAGATATTTGAAAC	263	
Db	194	AAATGCAACATGCGACGAGCAATGGAAAACAGAACAGCTGGGTGTTGAGATATTTGAAAC	253	
QY	264	TGCGGACTGTGAGGAGAATATTGAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATGT	323	
Db	254	TGCGGACTGTGAGGAGAATATTGAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATGT	313	
QY	324	TGACGGGTATTCGGAGTCAGCTTAAAAAGCTGCTTGCCGATACCGAAAAATATCATGG	383	
Db	314	TGACGGGTATTCGGAGTCAGCTTAAAAAGCTGCTTGCCGATACCGAAAAATATCATGG	373	
QY	384	CTACATGATGGCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGG	443	
Db	374	CTACATGATGGCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGG	433	
QY	444	ACCTCATTCAGACAGAAATCTATTACTTTGCCATGTCGTGTGAGACACAGAGAAATACACT	503	
Db	434	ACCTCATTCAGACAGAAATCTATTACTTTGCCATGTCGTGTGAGACACAGAGAAATACACT	493	
QY	504	GTTTATTCTGAAATTTCCCAAACTCATCAATAGACGACAGCTCTTAATGCTCTCTGGAA	563	
Db	494	GTTTATTCTGAAATTTCCCAAACTCATCAATAGACGACAGCTCTTAATGCTCTCTGGAA	553	
QY	564	GCCTCTTTTGGATCTTTTTCAGGGAACACTGGACTATGGAATGATTTCTCGAGAAGAAG	622	
Db	554	GCCTCTTTTGGATCTTTTTCAGGGAACACTGGACTATGGAATGATTTCTCGAGAAGAAG	612	

RESULT	72
AF173382	
LOCUS	
DEFINITION	Homo sapiens MSTP097 (MST097) mRNA linear
ACCESSION	AF173382
VERSION	AF173382.1
KEYWORDS	HTC; GI:33338001
SOURCE	Homo sapiens (human)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	HTC 01-AUG-2003

REFERENCE	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1. (bases 1 to 1278) Xu, H. S. Y., Zhao, B., Tong, Y. K., Sheng, H., Qin, B. M., Liu, Y. Q., Liu, B., Wang, A. Y., Zhang, Q., Song, L., Gao, Y., Zhang, C. L., Ye, J., Ji, X. J., Liu, B. H., Lu, H., Chen, J. Z., Cai, M. Q., Zheng, W. Y., Teng, C. Y., Liu, Q., Xu, L. T., Lin, J., Gong, Q., Zhang, A. M., Gao, R. L. and Hiu, R. T.
TITLE	Homo sapiens normal aorta mRNA MST097
JOURNAL	Unpublished
REFERENCE	2. (bases 1 to 1278) Xu, H. S. Y., Zhao, B., Tong, Y. K., Sheng, H., Qin, B. M., Liu, Y. Q., Liu, B., Wang, A. Y., Zhang, Q., Song, L., Gao, Y., Zhang, C. L., Ye, J., Ji, X. J., Liu, B. H., Lu, H., Chen, J. Z., Cai, M. Q., Zheng, W. Y., Teng, C. Y., Liu, Q., Xu, L. T., Lin, J., Gong, Q., Zhang, A. M., Gao, R. L. and Hiu, R. T.
AUTHORS	Direct Submission
TITLE	Submitted (30-JUN-1999) Molecular Medicine Center for Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167
JOURNAL	Bei Li Shi Lu, Beijing 100037, P.R. China
FEATURES	Location/Qualifiers
source	1..1278 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /tissue_type="aorta"
gene	1..1278 gene="MST097"
CDS	591..1010 /gene="MST097" /codon_start=1 /product="MSTP097" /protein_id="AAQ13623.1" /db_xref="GI:33338002" /translation="MEIFSPDGI CRGQISSGAIAGPSHVSIMNQIHGYMGLTRM QGYHRSVPSQAEKPFPPFNRLILHLHGLSNVHFATISILLSFLVRACKPYDLQIYP QYHRSVPSQAEKPFPPFNRLILHLHGLSNVHFATISILLSFLVRACKPYDLQIYP

Query Match	17.5%	Score 547;	DB 6;	Length 1278;
Best Local Similarity	100.0%;	Prod. No. 5.4e-284;		
Matches 547;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1844	CTGAGAGGTGACAAGGCTCACTGACCGTGGGTACTCATATCTTTGCTGCATCAGTCAGC	1903		
1	CTGAGAGGTGACAAGGCTCACTGACCGTGGGTACTCATATCTTTGCTGCATCAGTCAGC	60		
1904	ACTGTGACTCTTTATAAGTAAGTATAGTAACACAGAAGAAATCCACACATGTGTGTCCTTTT	1963		
61	ACTGTGACTCTTTATAAGTAAGTATAGTAACACAGAAGAAATCCACACATGTGTGTCCTTTT	120		
1964	ACRAGCTATCAAGTCTCTGAAGATGACCCAACTTGC AAAACAAAGAAATTTTGGGCGCACCA	2023		
121	ACRAGCTATCAAGTCTCTGAAGATGACCCAACTTGC AAAACAAAGAAATTTTGGGCGCACCA	180		
2024	TTTTGGATTCAGCAGGTCTCTTCCTGACTATACTCTCCAGAAATTTTCTCTTTTGAAA	2083		
181	TTTTGGATTCAGCAGGTCTCTTCCTGACTATACTCTCCAGAAATTTTCTCTTTTGAAA	240		
2084	GTACTACTGGATTACATGTATGGGATGCTCTCAAGCCCTCATGATCTACAGGCTCGAA	2143		
241	GTACTACTGGATTACATGTATGGGATGCTCTCAAGCCCTCATGATCTACAGGCTCGAA	300		
2144	AGAAATATCCTATCTGCTGTTTCATPATATGTTGTCCTCAGGTGCAGTTCGGTGAATAATC	2203		
301	AGAAATATCCTACTGTGCTGTTTCATPATATGTTGTCCTCAGGTGCAGTTCGGTGAATAATC	360		
2204	GGTTTAAAGAGTCAAGTATTTCCGTTTGAATACCCCTAGCCTCTCTAGGTTATGTGGTTG	2263		
361	GGTTTAAAGAGTCAAGTATTTCCGTTTGAATACCCCTAGCCTCTCTAGGTTATGTGGTTG	420		
2264	TAGTGNATAGACAAAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGCGCCCTTTAAAT	2323		
421	TAGTGNATAGACAAAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGCGCCCTTTAAAT	480		

(University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN	Query Match	Best Local Similarity	17.4%; Score 543; DB 3; Length 897;	Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1635	ATATAAAGCAGTCCAGTGGTGGCTGCCTGCTCCAGTGAATTCAGGTGCTCTATCAAGA	1694	
Db	1	ATATAAAGCAGTCCAGTGGTGGCTGCCTGCTCCAGTGAATTCAGGTGCTCTATCAAGA	60	
Qy	1695	GGAGATAGCAATACCAAGTGGTGAATGGGAAGTTCCTGGCGGATGGATCTTAATATCA	1754	
Db	61	GGAGATAGCAATACCAAGTGGTGAATGGGAAGTTCCTGGCGGATGGATCTTAATATCA	120	
Qy	1755	AGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCA	1814	
Db	121	AGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCA	180	
Qy	1815	TCACCTGTACGTAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT	1874	
Db	181	TCACCTGTACGTAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT	240	
Qy	1875	CTACTCACATCTTGTGCTGATCAGTACGTACGTACGTACGTACGTACGTACGTACGTACGT	1934	
Db	241	CTACTCACATCTTGTGCTGATCAGTACGTACGTACGTACGTACGTACGTACGTACGTACGT	300	
Qy	1935	CCAGAAAGATCCACACTGTGTGCTCTTTTCAAGCTATCAAGTCTTGAAGATGACCCCAAC	1994	
Db	301	CCAGAAAGATCCACACTGTGTGCTCTTTTCAAGCTATCAAGTCTTGAAGATGACCCCAAC	360	
Qy	1995	TTGCAAAACAAAGAAATTTGGCCACCAATTTTGGATTGAGTTCAGTCTTCTTCTGACTA	2054	
Db	361	TTGCAAAACAAAGAAATTTGGCCACCAATTTTGGATTGAGTTCAGTCTTCTTCTGACTA	420	
Qy	2055	TACTCTCCAGAAATTTCTTTTGAAGTACTACTGATTTTGAAGTACTACTGATTTTGAAGT	2114	
Db	421	TACTCTCCAGAAATTTCTTTTGAAGTACTACTGATTTTGAAGTACTACTGATTTTGAAGT	480	
Qy	2115	CTACAGCTCTATGATCTACAGCTTGGAAAGAAATATCTACTGTGCTGTTTATATATGG	2174	
Db	481	CTACAGCTCTATGATCTACAGCTTGGAAAGAAATATCTACTGTGCTGTTTATATATGG	540	
Qy	2175	TGG 2177		
Db	541	TGG 543		

RESULT 75
 DA809185
 LOCUS DA809185 OCBFP3 Homo sapiens 541 bp mRNA linear EST 11-NOV-2005
 DEFINITION DA809185 OCBFP3 Homo sapiens cDNA clone OCBFP3025332 5', mRNA
 ACCESSION DA809185
 VERSION DA809185.1 GI:82115632
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 541)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Teurittani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushi, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes

JOURNAL
 PUBMED
 COMMENT

Genome Res. 16 (1), 55-65 (2006)
 16344560
 Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Komatsuri, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3966
 Email: flj_cdn@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction; Helix Research Institute (HRI); 5'-end one pass sequencing; HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES

Location/Qualifiers
 source
 1..541
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OCBFP3025332"
 /tissue_type="brain"
 /dev_stage="fetal"
 /clone_lib="OCBFP3"
 /note="Vector: pME18SFL3"

ORIGIN

Query Match	Best Local Similarity	17.3%; Score 541; DB 9; Length 541;	Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	2	AGTGCTAAAGCCCTCCGAGGCCAAGCGCTGCTACTTGGCGCGCTCTTCTTAGTGGCGC	61
Db	1	AGTGCTAAAGCCCTCCGAGGCCAAGCGCTGCTACTTGGCGCGCTCTTCTTAGTGGCGC	60
Qy	62	GTTGGCGCGCTGGTGTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	121
Db	61	GTTGGCGCGCTGGTGTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120
Qy	122	AGTGGAGGGCG	181
Db	121	AGTGGAGGGCG	180
Qy	182	CCGGCG	241
Db	181	CCGGCG	240
Qy	242	TGGGTGTTGAGATATTTGAACTCGGACTGTGAGGAGATATTTGAATCAAGGATCGGC	301
Db	241	TGGGTGTTGAGATATTTGAACTCGGACTGTGAGGAGATATTTGAATCAAGGATCGGC	300
Qy	302	CTAAATTTGAGCCTTTTATGTTTCAGCGGTATTTCTTGGAGTCAAGTAAAAAGCTGCTTG	361
Db	301	CTAAATTTGAGCCTTTTATGTTTCAGCGGTATTTCTTGGAGTCAAGTAAAAAGCTGCTTG	360
Qy	362	CCGATACCAGAAATATCATGATGCTTAAGGCACCATGATTTTCATGTTTG	421
Db	361	CCGATACCAGAAATATCATGATGCTTAAGGCACCATGATTTTCATGTTTG	420
Qy	422	TGAAGAGATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCTG	481
Db	421	TGAAGAGATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCTG	480
Qy	482	GTGAGACACAGAAATATCATGCTTTTATTTCTGAAATCTCCCAAACTCAATAGAGCAG	541
Db	481	GTGAGACACAGAAATATCATGCTTTTATTTCTGAAATCTCCCAAACTCAATAGAGCAG	540
Qy	542	C 542	
Db	541	C 541	

DB339434/C
LOCUS DB339434 TEST14 Homo sapiens cDNA clone linear EST 04-DEC-2005
DEFINITION DB339434 sequence.
ACCESSION DB339434
VERSION DB339434.1 GI:83174333
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushiida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S. Diversification of transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan): cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES
source Location/Qualifiers
1..556
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TEST14003235"
/tissue type="testis"
/clone_lib="TEST14"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 17.3%; Score 541; DB 9; Length 556;
Best Local Similarity 100.0%; Pred. No. 9.2e-281; Indels 0; Gaps 0;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2559 CCCTGACGAGATGAACAGGCGTATTACTTAGGATCTGTGGCCATCAAGCAGAAAGTT 2618
DB 541 CCCTGACGAGATGAACAGGCGTATTACTTAGGATCTGTGGCCATCAAGCAGAAAGTT 482
QY 2619 CCCCTCTGAACCAATCGTTTACTGTCTTACATGGTTTCTCGATGAGATGTCATT 2678
DB 481 CCCCTCTGAACCAATCGTTTACTGTCTTACATGGTTTCTCGATGAGATGTCATT 422
QY 2679 TGCACATACCGATATTACTGAGTTTCTAGTGAGGCGCTGGAAGCCATATGATTACA 2738
DB 421 TGCACATACCGATATTACTGAGTTTCTAGTGAGGCGCTGGAAGCCATATGATTACA 362
QY 2739 GATCTATCCTCAGGAGAGACACAGCATAGAGTTTCTGATCGGGAGCAATATCACT 2798
DB 361 GATCTATCCTCAGGAGAGACACAGCATAGAGTTTCTGATCGGGAGCAATATGACT 302
QY 2799 GCATCTTTTGCATCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTTAAAGTGAT 2858
DB 301 GCATCTTTTGCATCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTTAAAGTGAT 242
QY 2859 ATAATTTTGACCTGTGTAGAACTCTCTGATATACACTGGCTATTATTAACCAATGAGGAGG 2918

DB 241 ATAATTTTGACCTGTGTAGAACTCTCTGATATACACTGGCTATTATTAACCAATGAGGAGG 182
QY 2919 TTTAATCAACAGAAAACACAGAAATTGATCATCATCATTTTGTGATACCTGCCATGTAACATCT 2978
DB 181 TTTAATCAACAGAAAACACAGAAATTGATCATCATCATTTTGTGATACCTGCCATGTAACATCT 122
QY 2979 ACTCTGAAAAATAATGTGTGCCATGCGAGGGCTCTACGGTTTGTGTAGTAACTTAATA 3038
DB 121 ACTCTGAAAAATAATGTGTGCCATGCGAGGGCTCTACGGTTTGTGTAGTAACTTAATA 62
QY 3039 CTTTAAACCCACATGCTCAAAATCAATGATACATATTTCTGAGAGACCCAGCAATACCA 3098
DB 61 CTTTAAACCCACATGCTCAAAATCAATGATACATATTTCTGAGAGACCCAGCAATACCA 2
QY 3099 T 3099 540 bp mRNA linear EST 06-NOV-2005
DB 1 T 1 CTONG2 Homo sapiens cDNA clone CTONG2004782 5', mRNA
DEFINITION DA437440 sequence.
LOCUS DA437440
DEFINITION DA437440.1 GI:80893993
ACCESSION DA437440
VERSION DA437440
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 540)
Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushiida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S. Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan): cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES
source Location/Qualifiers
1..540
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2004782"
/tissue type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 17.3%; Score 539; DB 9; Length 540;
Best Local Similarity 100.0%; Pred. No. 1.1e-279; Indels 0; Gaps 0;
Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Thu Jun 22 09:04:26 2006

Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1..569
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TRACH2007313"
/tissue_type="trachea"
/clone_lib="TRACH2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 17.3%; Score 539; DB 9; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.1e-279;
Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGTGTCTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCTCGCGCTCTTCTTAGTGCCG 60
DB 1 AAGTGTCTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCTCGCGCTCTTCTTAGTGCCG 60
QY 61 CGTTGCGCGCTGGTTGTACCGCGCGCGCGCGCGAGGAGCCACTGCAACCGAGACCG 120
DB 61 CGTTGCGCGCTGGTTGTACCGCGCGCGCGCGCGAGGAGCCACTGCAACCGAGACCG 120
QY 121 GAGTGGAGCGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGACGG 180
DB 121 GAGTGGAGCGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGACGG 180
QY 181 TCCGGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
DB 181 TCCGGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY 241 CTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGATATTTGAACTGCGGATCGG 300
DB 241 CTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGATATTTGAACTGCGGATCGG 300
QY 301 CCTAAATTGGAGCCTTTTATTTGAGCGGTATTTCTGAGTCACTGAGTCACTGAGTCACT 360
DB 301 CCTAAATTGGAGCCTTTTATTTGAGCGGTATTTCTGAGTCACTGAGTCACTGAGTCACT 360
QY 361 GCGATACAGAGAAATATCATGGCTACATGATGCTAAGGACCAATGATGATGATGATGAT 420
DB 361 GCGATACAGAGAAATATCATGGCTACATGATGCTAAGGACCAATGATGATGATGATGAT 420
QY 421 GTGAAGAGGAATGATCAGATGAGCTCATTCAGACAGATCTATTACCTTGCATGTCT 480
DB 421 GTGAAGAGGAATGATCAGATGAGCTCATTCAGACAGATCTATTACCTTGCATGTCT 480
QY 481 GGTGAGNACAGAGAAATACACTGTTTATTTCTGAAATTCGAAATTCGAAATTCGAAAT 539
DB 481 GGTGAGNACAGAGAAATACACTGTTTATTTCTGAAATTCGAAATTCGAAATTCGAAAT 539

RESULT 80
DA446578

LOCUS DA446578 CTONG2 Homo sapiens cDNA clone CTONG2017644 5', mRNA
DEFINITION sequence.
ACCESSION DA446578
VERSION DA446578.1 GI:80962067
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 591)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakeduri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushiida, N., M.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1..591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2017644"
/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="Vector: pME18SFL3"

FEATURES
source

ORIGIN

Query Match 17.3%; Score 539; DB 9; Length 591;
Best Local Similarity 99.8%; Pred. No. 1.1e-279;
Matches 589; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 506 TTTATTTCTGAATTTCCCAAACTATCAATAGACGACGAGTCTTAATGCTCTCTTGGAGC 565
DB 2 TTTATTTCTGAATTTCCCAAACTATCAATAGACGACGAGTCTTAATGCTCTCTTGGAGC 61
QY 566 CTCCTTTGGATCTTTTTCAGGCAACCTGACCTATGGAATGATTTCTCGAGAAGAAGAAC 625
DB 62 CTCCTTTGGATCTTTTTCAGGCAACCTGACCTATGGAATGATTTCTCGAGAAGAAGAAC 121
QY 626 TATTAAGAGAGAAGAAAACCGATTCGAAACAGTCGGAATTCCTTACGATTAATCAACAAG 685
DB 122 TATTAAGAGAGAAGAAAACCGATTCGAAACAGTCGGAATTCCTTACGATTAATCAACAAG 181
QY 686 GAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCG 745
DB 182 GAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCG 241
QY 746 CACAAGGATTTACGCAACAACTTTAAGCCCAATCTAGTGAAATCTAGTTGTCCTCCCAACA 805
DB 242 CACAAGGATTTACGCAACAACTTTAAGCCCAATCTAGTGAAATCTAGTTGTCCTCCCAACA 301
QY 806 TACGGATGGATCCAAAATTTATGCCCGCTGTATCCAGACTGGATTGCTTTTATACATAGCA 865

Db	302	TACGGATGGATCCAAAATTATGCTGCTGATCCAGACTGGATGCTTTTATACATAGCA	361
Qy	866	ACGATATTTGGATATCTAAACATCGTACCCAGAGAGAAAGGAGACTCACTTATGTGCACA	925
Db	362	ACGATATTTGGATATCTAAACATCGTACCCAGAGAGAAAGGAGACTCACTTATGTGCACA	421
Qy	926	ATGAGCTAGCCAAACATCGAGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTGTGTTCTCC	985
Db	422	ATGAGCTAGCCAAACATCGAGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTGTGTTCTCC	481
Qy	986	AAGAAGATTTGATAGATATTCCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAATCCCA	1045
Db	482	AAGAAGATTTGATAGATATTCCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAATCCCA	541
Qy	1046	GTGGTGTGTAATAATTCCTAGAAATCTATATGAAGAAAATGATGAATCTGAG	1095
Db	542	GTGGTGTGTAATAATTCCTAGAAATCTATATGAAGAAAATGATGAATCTGAG	591

```

RESULT_81
BU784521/c
LOCUS
DEFINITION
3', similar to TR:075273 R25984_1 ; mRNA sequence.
EST 11-OCT-2002
IMAGE:6124487
BU784521
597 bp mRNA linear
BU784521
Accession
BU784521
Version
BU784521.1
Keywords
GI:23829606
Source
Homo sapiens (human)
Organism
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 597)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Secorce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hiller, L., Marra, M., Page, D., Wylie, T., Martin, J., Bliscain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, J., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, E., Cole, R., Tsagarieishvili, R.,
Williams, T., Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557

```

8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library.

Query Match	17.2%; Score 538; DB 3; Length 597;	
Best Local Similarity	99.8%; Pred. No. 3.9e-279;	
Matches 588; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	2532 AGGATACACGGAACGTTATATGGTCAACCTGACAGATGACAGGCTATTACTTAGG	2591
DB		
QY	593 AGGATACACGGAACGTTATATGGTCAACCTGACAGATGACAGGCTATTACTTAGG	534
DB		
QY	2592 ATCTGTGCCCATGCAACGAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACA	2651
DB		
QY	533 ATCTGTGCCCATGCAACGAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACA	474
DB		
QY	2652 TGGTTTCCTGGATGAGAAATGTCATTTTGCAATACACAGTATATTACTGAGTTTTTAACT	2711
DB		
QY	473 TGGTTTCCTGGATGAGAAATGTCATTTTGCAATACACAGTATATTACTGAGTTTTTAACT	414
DB		
QY	2712 GAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGCT	2771
DB		
QY	413 GAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGCT	354
DB		
QY	2772 TCCTGAATCGGAGAAACATTATGAACCTGCATCTTTTGCACCTACTCTTCAAGAAAACCTTGG	2831
DB		
QY	353 TCCTGAATCTGGAGAAACATTATGAACCTGCATCTTTTGCACCTACTCTTCAAGAAAACCTTGG	294
DB		
QY	2832 ATCAACGATTTGCTGCTCTAAAGTGTATATAATTTTGACCTGTGTAGAACTCTCTGGTATA	2891
DB		
QY	293 ATCAACGATTTGCTGCTCTAAAGTGTATATAATTTTGACCTGTGTAGAACTCTCTGGTATA	234
DB		
QY	2892 CACTGCTATTTTAAACCAATCAGGAGGTTTAAATCAACGAAACACACAGATTTGATCATCA	2951
DB		
QY	233 CACTGCTATTTTAAACCAATCAGGAGGTTTAAATCAACGAAACACACAGATTTGATCATCA	174
DB		
QY	2952 CATTTTGTATACCTGCATGTAAACATCTACTCTGGAATAATTTGGTGCCATGCAGGG	3011
DB		
QY	173 CATTTTGTATACCTGCATGTAAACATCTACTCTGGAATAATTTGGTGCCATGCAGGG	114
DB		
QY	3012 TCTACGGTTTGGTGTAGTAACTTAATACCTTAAACCCCATGCTCAAAATCAAAATGATAC	3071
DB		
QY	113 TCTACGGTTTGGTGTAGTAACTTAATACCTTAAACCCCATGCTCAAAATCAAAATGATAC	54
DB		
QY	3072 ATATTCTCTGAGAGCCAGCATACCATAGAAATTCTAAAAAATAAAAA 3120	
DB		
QY	53 ATATTCTCTGAGAGCCAGCATACCATAGAAATTCTAAAAAATAAAAA 5	
DB		

was constructed by Dr. J. Ferrer In vivo mass-excised to
 pBluecript Sk- by Dr. H. Inoue DNA sequencing by: Washington
 University Genome Sequencing Center for information on obtaining a
 clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
 Seq primer: 4096 from Gibco
 High quality sequence stop: 490.
 Location/Qualifiers
 1..597
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

FEATURES
 source

RESULT 82
 DAS15446
 CUS
 DEFINITION

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Db 361 CATGTTTGTGAAGAGGAATGATCCAGATGGAGCTCATTCAGACAGAAATCTATTACCTTGC 420
 QY 474 CATGTTCTGTGTGAGAACAGAGAAAATACACTGTGTTTATTCTGAAATTCCTCCAAACTATCAA 533
 Db 421 CATGTTCTGTGTGAGAACAGAGAAAATACACTGTGTTTATTCTGAAATTCCTCCAAACTATCAA 480
 QY 534 TAGAGCAGCAGTCTTAATGCTCTCTTGTGAAGCCCTCTTTTGGATCTTTTTCAGGCAACT 593
 Db 481 TAGAGCAGCAGTCTTAATGCTCTCTTGTGAAGCCCTCTTTTGGATCTTTTTCAGGCAACT 540
 QY 594 GGAATATGAAATGATTTCTTCAGAGAGAGAACTATTAAAGAGAAAG 638
 Db 541 GGAATATGAAATGATTTCTTCAGAGAGAGAACTATTAAAGAGAAAG 585

RESULT 84
 DA773829
 LOCUS
 DEFINITION
 DA773829 OCBPF2 Homo sapiens cDNA clone OCBF2007376 5', mRNA
 VERSION
 DA773829
 KEYWORDS
 DA773829.1 GI:81475405
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM

REFERENCE
 AUTHORS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominiidae; Homo.

1 (bases 1 to 578)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaguri, H.,
 Iehii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kuehida, N.,
 Yonekawa, T., Otsuka, S., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
 Murakami, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Title
 Diversification of Transcriptional Modulation: Large-scale
 Identification and Characterization of Putative Alternative
 Promoters of Human Genes

JOURNAL
 PUBMED
 COMMENT
 Genome Res. 16 (1), 55-65 (2006)
 Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com

NESTO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction;
 Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation, 3'-end one
 pass sequencing; RAB.

FEATURES
 source
 Location/Qualifiers
 1..578
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OCBPF2007376"
 /tissue_type="brain"
 /dev_stage="fetal"
 /clone_lib="OCBPF2"
 /note="Vector: pME18SFL3"

ORIGIN
 Query Match 16.9%; Score 527; DB 9; Length 578;
 Best Local Similarity 99.8%; Pred. No. 3 6e-273;
 Matches 577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGCTACTCCGCCCGCTCTTCTTAGTGCCG 60
 Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGCTACTCCGCCCGCTCTTCTTAGTGCCG 60

QY 61 CGTTCGCCCTGGGTTGTCAACCGCGCGCGCGGAGGAGCCACTGCAACCCAGGACCG 120
 Db 61 CGTTCGCCCTGGGTTGTCAACCGCGCGCGCGGAGGAGCCACTGCAACCCAGGACCG 120
 QY 121 GAGTGAGGCGCGCGAGCATGAGCGGCGGCGGCGGCTCATAGCGCACGTCGGGACGG 180
 Db 121 GAGTGAGGCGCGCGAGCATGAGCGGCGGCGGCGGCTCATAGCGCACGTCGGGACGG 180
 QY 181 TCCGCGCGCGCGCGGAGGAGAAATGCAATGCGCAGCAGCAATGGAACAGAAACAG 240
 Db 181 TCCGCGCGCGCGCGGAGGAGAAATGCAATGCGCAGCAGCAATGGAACAGAAACAG 240
 QY 241 CTGGGCTGTGAGATATTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
 Db 241 CTGGGCTGTGAGATATTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
 QY 301 CCTAANTTGGAGCCCTTTTATTGTGTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360
 Db 301 CCTAANTTGGAGCCCTTTTATTGTGTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360
 QY 361 GCCGATACCCAGAAATATCATGGCTACATGATGGCTTAAGGCACCATGATTTTCATGTTT 420
 Db 361 GCCGATACCCAGAAATATCATGGCTACATGATGGCTTAAGGCACCATGATTTTCATGTTT 420
 QY 421 GTGAGAGGATGATTCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCT 480
 Db 421 GTGAGAGGATGATTCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCT 480
 QY 481 GGTGAGAAACAGAGAAATACACTGTTTATTCTGAAATTTCCAAACTATCAATAGAGCA 540
 Db 481 GGTGAGAAACAGAGAAATACACTGTTTATTCTGAAATTTCCAAACTATCAATAGAGCA 540
 QY 541 GCAGTCTTAATGCTCTCTTGTGAAGCCTCTTTTGGATCT 578
 Db 541 GCAGTCTTAATGCTCTCTTGTGAAGCCTCTTTTGGATCT 578

RESULT 85
 BE888665
 LOCUS
 DEFINITION
 601513061F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3914582 5',
 mRNA sequence.

ACCESSION
 BE888665
 KEYWORDS
 BE888665.1 GI:10345197
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominiidae; Homo.

1 (bases 1 to 674)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbe@mail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 DNA distribution by: Incyte Genomics, Inc.
 Cloned through the MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M9736 Row: i column: 15
 High quality sequence stop: 662.
 Location/Qualifiers
 1..674
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3914582"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"

FEATURES
 source

```
/clone lib="NIH MGC 71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2.1 kb. "
```

16.9%; Score 526; DB 7; Length 674;
tv 100.0%; Pred. No. 1.3e-272;

GAGCATCACCTGTACGTAGTCAGTTACGTAAATCTGGAGAGGTGACAGGCT 60

CGTGGCTACTCACAATCTTTGCTGCATCAGTCAGCACTGTGACATCTTTATAAG 1922
CGTGGCTACTCACAATCTTTGCTGCATCAGTCAGCACTGTGACATCTTTATAAG 120
AGTAACCGAAGAANTCCACACTGTGTGTCCCTTTTACAAGCTATCAAGTCCTCTGA 1982
AGTAACCGAAGAANTCCACACTGTGTGTGCCCTTTTACAAGCTATCAAGTCCTCTGA 180

CCAACTTGCAAAACAAAGGAATTTGGGCCACCAATTTGGAATTCAGCAGGTCC 2042

CCCAACTTGCAAAACAAAGGAATTTGGGCCACCATTTTGGATTTCAGCAGGTCC 240

GACTATACTCCTCCAGAAATTTTCTCTTTTTTGAAAGTACTACTGGATTTCACAT 2102

CGACTATACTCCTCCAGAAATTTTCTCTTTTGAAAGTACTACTGGATTACATTT 300

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

5'ATGTGCTCTACAGGCTCATGATCTACAGCCCTGGATAGAGAAATATCCATACCTGTC1303

ATATGGTGGGTCCTCAGGTGCCAGTGGTGAAATACCGTTAATCGGTTAAAGGAGTCAAGTA 420

ATATGGGCGGCTTCAGGTCAGTGGTGGTGTAGTATAGACACAGGGG 2282

CTTGGAATACCCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGG 480

TCACCGAGGGCTTAAATTTGAAGGCGCCTTTAAATATAA 2328

TTACCGAGGGCTTAAATTTGAAGGGCGCTTTTAAATATAAA 526

EST 05-NOV-2005

NT2NE2	Homo sapiens	cdna clone	NT2NE2014882	5', mRNA
		565 bp	mRNA	linear

00000000

.I G1:80//2330
ions (human)

iens (human)
iens
iens
a. Metazoa: Chordata: Craniata: Vertebrata; Euteleostomi;
iens

a, Metazoa; Eutheria; Euarchontoglires; Primates; Catarrhini; e: Homo.

Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,

a, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,

T. Otsuka, R. Kanda, K. Yokoi, T. Kondo, H. Wagatsuma, M. Ishibashi, T. Takahashi-Fujii, A. Ishida, S. Sugano

Terada, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, I. and Sugano, S. (1991) Regulation of Transcriptional Modulation: Large-scale characterization of putative alternative promoters in the human genome. *Proc. Natl. Acad. Sci. USA* 88: 1053-1057.

Publication and Characterization of Putative Alternative
 splicing of Human Genes
 10-11-01 11:55:55 (2006)

Res. 16 (1), 55-65 (2006)

1. Takao Iwogai

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1717)
AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Kueguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kuehida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isegai, T. and Sugano, S.

TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes

JOURNAL Promoters of Human Genes
PUBLISHED Genome Res. 16 (1), 55-65 (2006)
COMMENT 1634560
Contact: Takao Isegai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute (HRI); cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and HRI.

FEATURES
source Location/Qualifiers
1..717
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLAC81002059"
/tissue_type="placenta"
/clone_lib="PLAC81"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 16.8%; Score 525; DB 9; Length 717;
Best Local Similarity 100.0%; Pred. No. 4,4e-272; Indels 0; Gaps 0;
Matches 525; Conservative 0; Mismatches 0;

QY 1482 TGAAGAACACACACATCTGGATTAATATCCATGACATCTTTCATGTTTCCCAAG 1541
DB 1 TGAAGAACACACACATCTGGATTAATATCCATGACATCTTTCATGTTTCCCAAG 60

QY 1542 TCACGAGAGGAATTCAGTTTATTTTCCTCTGATGCAAAACAGGTTTCCGTCAATT 1601
DB 61 TCACGAGAGGAATTCAGTTTATTTTCCTCTGATGCAAAACAGGTTTCCGTCAATT 120

QY 1602 ATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGCATCCAGTGGTGGCTGCC 1661
DB 121 ATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGCATCCAGTGGTGGCTGCC 180

QY 1662 TGCTCCAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATG 1721
DB 181 TGCTCCAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATG 240

QY 1722 GGAAGTTCTTGGCGGCGATGGATCTAATATCAAGTTGATGAAGTCAAGAGGCTGGGTATA 1781
DB 241 GGAAGTTCTTGGCGGCGATGGATCTAATATCAAGTTGATGAAGTCAAGAGGCTGGGTATA 300

QY 1782 TTTTGAAGGCACAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTGAA 1841
DB 301 TTTTGAAGGCACAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTGAA 360

QY 1842 TCCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCACTCA 1901
DB 361 TCCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCACTCA 420

QY 1902 GCACGTGTGACTTCTTTTAAAGTAAGTAGTAGTAACCAAGAAAGATCCACACTGTGTCCCT 1961

Db 421 GCACGTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCT 480

QY 1962 TTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAA 2006

Db 481 TTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAA 525

RESULT 88
DR002989
LOCUS
DEFINITION
DR002989 Human fetal brain, large insert, pcMV expression library
TCl25354
Homo sapiens cDNA clone TCl25354, 5', similar to Homo sapiens dipeptidylpeptidase 8 (DPP8), transcript variant 2, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

DR002989 523 bp mRNA linear EST 17-MAY-2005
TCl25354 Human fetal brain, large insert, pcMV expression library
Homo sapiens cDNA clone TCl25354, 5', similar to Homo sapiens dipeptidylpeptidase 8 (DPP8), transcript variant 2, mRNA sequence.
DR002989
DR002989.1 GI:66262862
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo
1 (bases 1 to 523)
Birkett, C., Cho, J., Gau, Y., Hamer, R., Kelly, S., Kovacs, K., Liu, L., Liu, X., Porter, J., Sachs, A., Shu, Y., Sun, Z., Wong, J., Wu, M., Zhang, X., Jay, G. and He, W.
High-throughput cloning of full-length human cDNAs directly from cDNA libraries optimized for large and rare transcripts
Unpublished (2005)
Contact: Kovacs, KF
High Throughput cDNA Cloning
Origene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cdna@origene.com
This EST submission is part of an on-going human full-length cloning project at Origene Technologies, Inc.
Please contact Origene for access.
Origene Technologies, Inc.
6 Taft Ct, Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
http://www.origene.com
Seq primer: pCMV6 5prime forward vector primer, Origene Technologies Inc.
Location/Qualifiers
1..523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TCl25354"
/tissue_type="Fetal Brain"
/clone_lib="Human fetal brain, large insert, pcMV expression library"
/notes="Organ: Fetal Brain; Vector: pCMV6-XL4; Site 1: EcoRI; Site 2: XhoI/SalI compatible end ligatio; Oligo-dt primed reverse transcription optimized for large and gc rich mRNA transcripts, cDNA size selection, optimized ligation for large inserts into mammalian expression vector, random clones selected for end sequence verification of full-length genes"

FEATURES
source Location/Qualifiers
1..523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TCl25354"
/tissue_type="Fetal Brain"
/clone_lib="Human fetal brain, large insert, pcMV expression library"
/notes="Organ: Fetal Brain; Vector: pCMV6-XL4; Site 1: EcoRI; Site 2: XhoI/SalI compatible end ligatio; Oligo-dt primed reverse transcription optimized for large and gc rich mRNA transcripts, cDNA size selection, optimized ligation for large inserts into mammalian expression vector, random clones selected for end sequence verification of full-length genes"

ORIGIN
Query Match 16.8%; Score 523; DB 9; Length 523;
Best Local Similarity 100.0%; Pred. No. 5.3e-271; Indels 0; Gaps 0;
Matches 523; Conservative 0; Mismatches 0;

QY 20 GCCAAGCGCGCTCTACTCCCGCTGCTTCTTAGTCCCGCTTCGCGCTGGTGT 79
DB 1 GCCAAGCGCGCTCTACTCCCGCTGCTTCTTAGTCCCGCTTCGCGCTGGTGT 60

QY 80 CACCGGCGCGCGCGCGAGAACCACTGCAACCAAGGACCGAGTGGAGCGCGCAGCA 139

Thu Jun 22 09:04:26 2006

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="THYMU3021319"
/tissue_type="thymus"
/clone_lib="THYMU3"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 16.7%; Score 522; DB 9; Length 546;
Best Local Similarity 100.0%; Pred. No. 1.8e-270; Indels 0; Gaps 0;
Matches 522; Conservative 0; Mismatches 0

Qy 1 AAGTGTAAAGCTCCGAGGCAAGCGCGTGTACTGCGCGCGCTCTTCTAGTGGCG 60
Db 3 AAGTGTAAAGCTCCGAGGCAAGCGCGTGTACTGCGCGCGCTCTTCTAGTGGCG 62
Qy 61 CGTTCCGCGCGTGTGTACCGCGCGCGCGCGAGGAGCAATGCAACACGAGGACCG 120
Db 63 CGTTCCGCGCGTGTGTACCGCGCGCGCGCGAGGAGCAATGCAACACGAGGACCG 122
Qy 121 GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGAGGAGCAATGCAACACGAGGACCG 180
Db 123 GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGAGGAGCAATGCAACACGAGGACCG 182
Qy 181 TCCGCGCGCGCGCGGAGGAGCAATGCAACACGAGGAGCAATGCAACACGAGGACCG 240
Db 183 TCCGCGCGCGCGCGGAGGAGCAATGCAACACGAGGAGCAATGCAACACGAGGACCG 242
Qy 241 CTGGGTGTGAGATATTGAAACTGCGGACTGTGAGGAGCAATGCAACACGAGGACCG 300
Db 243 CTGGGTGTGAGATATTGAAACTGCGGACTGTGAGGAGCAATGCAACACGAGGACCG 302
Qy 301 CCTAAATTGAGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360
Db 303 CCTAAATTGAGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 362
Qy 361 GCGGATACAGAAATATCATGCTTACATGATGCTTAAGGACCAACATGATTTTCATGTTT 420
Db 363 GCGGATACAGAAATATCATGCTTACATGATGCTTAAGGACCAACATGATTTTCATGTTT 422
Qy 421 GTCAAGAGCAATGATCCAGATGAGCCTCATTCAGACAGCAATCTTACCTTGCATGCTCT 480
Db 423 GTCAAGAGCAATGATCCAGATGAGCCTCATTCAGACAGCAATCTTACCTTGCATGCTCT 482
Qy 481 GGTGAGAACAGAAATATCATGCTTACATGCTTAAATTTCC 522
Db 483 GGTGAGAACAGAAATATCATGCTTAAATTTCC 524

RESULT 90
DA990713 SYN04 Homo sapiens cDNA clone SYN04002393 5', mRNA
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

Db 61 CACCGCGCGCGCGCGAGGAGCCATGCAACGAGCGGAGTGGAGCGCGCGCAGCA 120
Qy 140 TGAAGCG 199
Db 121 TGAAGCG 180
Qy 200 AAGGAAATGCAACATGCGAGCAATGGAACAGCAACAGCTGGGTGTGAGATATTG 259
Db 181 AAGGAAATGCAACATGCGAGCAATGGAACAGCAACAGCTGGGTGTGAGATATTG 240
Qy 260 AACTCGCGAGCTGTGAGAGCAATTTGAATCAAGGATCGGCTTAAATTTGAGCGCTTTT 319
Db 241 AACTCGCGAGCTGTGAGAGCAATTTGAATCAAGGATCGGCTTAAATTTGAGCGCTTTT 300
Qy 320 ATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACAGAAATATC 379
Db 301 ATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACAGAAATATC 360
Qy 380 ATGCTACATGATGCTTAAGGACCAACATGATTTTCATGTTGAGAGCAATGATCCAG 439
Db 361 ATGCTACATGATGCTTAAGGACCAACATGATTTTCATGTTGAGAGCAATGATCCAG 420
Qy 440 ATGACCTCATTTACAGACAGCAATCTATACCTTGCATGCTGCTGAGACAGAGAAATA 499
Db 421 ATGACCTCATTTACAGACAGCAATCTATACCTTGCATGCTGCTGAGACAGAGAAATA 480
Qy 500 CACTGTTTATCTGAAATTTCCCAAACTATCAATAGACGAGC 542
Db 481 CACTGTTTATCTGAAATTTCCCAAACTATCAATAGACGAGC 523

RESULT 89
DB146190 THYMU3 Homo sapiens cDNA clone THYMU3021319 5', mRNA
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT
FEATURES
source

```

16344560
COMMENT Contact: Takao Isogai
          HRI Project (HRI Team)
          Helix Research Institute
          7-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
          Tel: 81-438-52-3975
          Fax: 81-438-52-3986
          Email: flj-cdna@nifty.com
          NEDO human cDNA project (New Energy and Industrial Technology
          Developmental Organization, Japan): cDNA library construction:
          Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
          Research Association for Biotechnology (RAB) and Biotechnology
          Center, National Institute of Technology and Evaluation; 3'-end one
          pass sequencing: RAB.
FEATURES   Location/Qualifiers
            1..523
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="SYNOV4002393"
            /tissue_type="synovial membrane tissue from rheumatoid
            arthritis"
            /clone_lib="SYNOV4"
            /note="Vector: pME18SFL3"
ORIGIN
Query Match      16.7%; Score 521; DB 9; Length 523;
Best Local Similarity 100.0%; Pred. No. 6.4e-270;
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGTGTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG 60
DB 3 AAGTGTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG 62
QY 61 CGTTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 63 CGTTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 122
QY 121 GAGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 123 GAGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 182
QY 181 TCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
DB 183 TCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 242
QY 241 CTGGGTGTTGAGATATTGAACTGCGGACTGTGAGGAGAAATATTGAATCACAGATCGG 300
DB 243 CTGGGTGTTGAGATATTGAACTGCGGACTGTGAGGAGAAATATTGAATCACAGATCGG 302
QY 301 CCTAAATTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360
DB 303 CCTAAATTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 362
QY 361 GCCGATACCGAATAATCATCGCTACATGATGGCTAAGGCACACATGATTTTCATGTTT 420
DB 363 GCCGATACCGAATAATCATCGCTACATGATGGCTAAGGCACACATGATTTTCATGTTT 422
QY 421 GTGAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTGCGCATGCT 480
DB 423 GTGAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTGCGCATGCT 482
QY 481 GGTGAGACAGAGAAATACACTGTGTTTATTCTGAAATTC 521
DB 483 GGTGAGACAGAGAAATACACTGTGTTTATTCTGAAATTC 523

RESULT 91
CN427232
LOCUS
DEFINITION 542 bp mRNA linear EST 16-MAY-2004
ACCESSION 17006060005171 GRN_PRENU Homo sapiens cDNA 5', mRNA sequence.
VERSION CN427232.1 GI: 47414826

```

```

KEYWORDS EST. Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
REFERENCE 1 (bases 1 to 542)
AUTHORS   Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
          Li, X., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
          Lebowski, J. and Stanton, L.W.
          Transcriptome characterization elucidates signaling networks that
          control human ES cell growth and differentiation
          Nat. Biotechnol. 22 (6), 707-716 (2004)
JOURNAL    15146197
PUBMED
COMMENT    Contact: Brandenberger R
          Regenerative Medicine
          Genon Corporation
          230 Constitution Drive, Menlo Park, CA 94025, USA
          Tel: 650 473 8658
          Fax: 650 473 7760
          Email: rbrandenberger@genon.com
          Insert Length: 542 Std Error: 0.00.
          Location/Qualifiers
            1..542
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="embryonic stem cell, retinoic acid and
            mitogen-treated hES cell line H7"
            /clone_lib="GRN_PRENU"
            /note="oligo dT primed, full-length enriched cDNA library
            from hES cell line H7 (p29) maintained in feeder-free
            conditions. Embryoid bodies were generated in the presence
            of all-trans retinoic acid and mitogens."
ORIGIN
Query Match      16.6%; Score 518; DB 8; Length 542;
Best Local Similarity 100.0%; Pred. No. 2.7e-268;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 68 CGCTTGGGTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 127
DB 1 CGCTTGGGTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 128 GCGCGCGCAGCATGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 187
DB 61 GCGCGCGCAGCATGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 188 GGGCGCGGGGGAAGGAAATGCAATGCGCAGCAGCAATGGAACAGAAACAGAGCTGGTG 247
DB 121 GGGCGCGGGGGAAGGAAATGCAATGCGCAGCAGCAATGGAACAGAAACAGAGCTGGTG 180
QY 248 TTGAGATATTGAACTGCGGACTGTGAGGAGAAATTTGAATCACAGATCGGCTAAAT 307
DB 181 TTGAGATATTGAACTGCGGACTGTGAGGAGAAATTTGAATCACAGATCGGCTAAAT 240
QY 308 TGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATA 367
DB 241 TGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATA 300
QY 368 CCAGAAATATCATCGCTACATGATGCTAAGGCACCAATGATTTTCATGTTTGAAGA 427
DB 301 CCAGAAATATCATCGCTACATGATGCTAAGGCACCAATGATTTTCATGTTTGAAGA 360
QY 428 GGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCTGGTGAGA 487
DB 361 GGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCTGGTGAGA 420
QY 488 ACAGAGGAATACACTGTTTTTATTCTGAAATTCCTGAAATTCCTGAAATTCCTGAAATTC 547
DB 421 ACAGAGGAATACACTGTTTTTATTCTGAAATTCCTGAAATTCCTGAAATTCCTGAAATTC 480
QY 548 TAATGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAG 585

```


Db 481 TTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAG 518

RESULT 92
LOCUS AW958665 639 bp mRNA linear EST 01-JUN-2000
DEFINITION EST370735 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION AW958665
VERSION AW958665.1 GI:8148349
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Hegde, P., Qi, R., Aherthy, K., Dharap, S., Gaspard, R., Gay, C., Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J. and Quackenbush, J.
AUTHORS Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
TITLE Unpublished (2000)
JOURNAL Contact: John Quackenbush
COMMENT The Institute for Genomic Research
712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: Johnq@tigr.org
Plate: 124
Seq primer: Reverse.
FEATURES
source Location/Qualifiers
1..639
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGE"
/note="Vector: pBluescriptSKm"

ORIGIN
Query Match 16.6%; Score 518; DB 7; Length 639;
Best Local Similarity 100.0%; Pred. No. 2.8e-268;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2560 CCTGACCAAGATGAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAAAAGTTC 2619
Db 518 CTGACCAAGATGAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAAAAGTTC 459
QY 2620 CCCTGTGAACCAATCGTTTACTGCTTACATGTTTCTCGATGAGATGCCATTTT 2679
Db 458 CCCTGTGAACCAATCGTTTACTGCTTACATGTTTCTCGATGAGATGCCATTTT 399
QY 2680 GCACATACCAGTATATTACTGATTTTTTTAGTGAGGCTGGAAGCCATATGATTACAG 2739
Db 398 GCACATACCAGTATATTACTGATTTTTTTAGTGAGGCTGGAAGCCATATGATTACAG 339
QY 2740 ATCTATCTCAGAGAGACACAGATAGAGTTCTGTAATCGGAGAACATATGAACTG 2799
Db 338 ATCTATCTCAGAGAGACACAGATAGAGTTCTGTAATCGGAGAACATATGAACTG 279
QY 2800 CATCTTTTGCACTACCTTCAAGAAAACCTTGATCAGCTATTGCTCTAAAGTGA 2859
Db 278 CATCTTTTGCACTACCTTCAAGAAAACCTTGATCAGCTATTGCTCTAAAGTGA 219
QY 2860 TAAATTTGACCTGTGTAGACTCTCTGTATATACCTGGCTATTTAAACAAATGAGAGGT 2919
Db 218 TAAATTTGACCTGTGTAGACTCTCTGTATATACCTGGCTATTTAAACAAATGAGAGGT 159
QY 2920 TTAATCAACAGAAAACACAGAAATGATCATCAATTTTGCATCTGCCATGTAACATCTA 2979
Db 158 TTAATCAACAGAAAACACAGAAATGATCATCAATTTTGCATCTGCCATGTAACATCTA 99
QY 2980 CTCCTGAAAATAAATGTGGTCCATGCAGGGGCTACGGGTTTGTGGTAGTAATCTAATAC 3039

Db 98 CTCCTGAAAATAAATGTGGTGGCCATGCAGGGGTACGGTTTGTGGTAGTAATCTAATAC 39

QY 3040 CTTAACCCCATGCTCTCAAAATCAAAATGATACATATTC 3077

Db 38 CTTAACCCCATGCTCTCAAAATCAAAATGATACATATTC 1

RESULT 93
LOCUS DA412358 517 bp mRNA linear EST 06-NOV-2005
DEFINITION DA412358 BRTHA3 Homo sapiens cDNA clone BRTHA3019884 5', mRNA sequence.
ACCESSION DA412358
VERSION DA412358.1 GI:80939509
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 517)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO Human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.
FEATURES
source Location/Qualifiers
1..517
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRTHA3019884"
/issue_type="thalamus"
/clone_lib="BRTHA3"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 16.6%; Score 517; DB 9; Length 517;
Best Local Similarity 100.0%; Pred. No. 9.5e-268;
Matches 517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GTGCTAAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTCTTCTTAGTCCGCG 62
Db 1 GTGCTAAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTCTTCTTAGTCCGCG 60
QY 63 TTCGCGCGCTGGTGTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122
Db 61 TTCGCGCGCTGGTGTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 123 GTGAGGCG 182
Db 121 GTGAGGCG 180

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT
CONTACT: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library constructing:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
source
1. .564
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NTONG2007406"
/tissue_type="tongue"
/clone_lib="NTONG2"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 16.4%; Score 513; DB 9; Length 564;
Best Local Similarity 99.8%; Pred. No. 1.4e-265;
Matches 563; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAGTCTAAAGCCCTCCGAGGCCAAGCGCGTGTCTACTGCGCGCGTCTTCTTAGTGGCG 60
DB 1 AAGTCTAAAGCCCTCCGAGGCCAAGCGCGTGTCTACTGCGCGCGTCTTCTTAGTGGCG 60
QY 61 CGTTCCGCGCTGGGTGTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 61 CGTTCCGCGCTGGGTGTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 GAGTGGAGCG 180
DB 121 GAGTGGAGCG 180
QY 181 TCCG 240
DB 181 TCCG 240
QY 241 CTGGGTGTTGAGATATTTGAAACTCGGACTGTGAGGAGATATTTGAATCAGAGATCGG 300
DB 241 CTGGGTGTTGAGATATTTGAAACTCGGACTGTGAGGAGATATTTGAATCAGAGATCGG 300
QY 301 CCTAAATTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGCTTAAAGCTGCTT 360
DB 301 CCTAAATTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGCTTAAAGCTGCTT 360
QY 361 GCCGATACCAAGAAAATATCATGCTTACATGCTTACGAGGACCAATGATTTTATGTTT 420
DB 361 GCCGATACCAAGAAAATATCATGCTTACATGCTTACGAGGACCAATGATTTTATGTTT 420
QY 421 GTGAAGAGGAATGATTCAGATGACCTTATTCAGACAGATATCTATTACCTTGGCATGCT 480
DB 421 GTGAAGAGGAATGATTCAGATGACCTTATTCAGACAGATATCTATTACCTTGGCATGCT 480

FEATURES
source
1. .658
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3613139"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC-44"
/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACAGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
Query Match 16.5%; Score 515; DB 7; Length 658;
Best Local Similarity 100.0%; Pred. No. 1.2e-266;
Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1811 AGCATCACTGTACGTAGTACGTTACGTTAAATCCTGGAGAGTGCACAGGCTGACTGACC 1870
DB 71 AGCATCACTGTACGTAGTACGTTAAATCCTGGAGAGTGCACAGGCTGACTGACC 130
QY 1871 GTGGCTACTCACATTTCTGCTGCATCAGTGCACACTGTGACTTCTTTATAGTAAGTATA 1930
DB 131 GTGGCTACTCACATTTCTGCTGCATCAGTGCACACTGTGACTTCTTTATAGTAAGTATA 190
QY 1931 GTACACAGAGATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCTGAATGACC 1990
DB 191 GTACACAGAGATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCTGAATGACC 250
QY 1991 CAATCTGCAAAACAAAGAAATTTTGGCCACCAATTTTGGATTCACAGCTCTCTCTCTG 2050
DB 251 CAATCTGCAAAACAAAGAAATTTTGGCCACCAATTTTGGATTCACAGCTCTCTCTCTG 310
QY 2051 ACTATACCTCTCCAGAAATTTCTTTTGAAGTACTACTGGAATTTACATTTATGGGA 2110
DB 311 ACTATACCTCTCCAGAAATTTCTTTTGAAGTACTACTGGAATTTACATTTATGGGA 370
QY 2111 TGCTCTACAAGCTCATGATCTACAGCTCGAAAGAAATTCCTACTGTGCTGTTTCATAT 2170
DB 371 TGCTCTACAAGCTCATGATCTACAGCTCGAAAGAAATTCCTACTGTGCTGTTTCATAT 430
QY 2171 ATGGTGGTCTCAGTGAGTGTGTAATTCGGTTTAAAGAGTCAAGTATTTCCGCT 2230
DB 431 ATGGTGGTCTCAGTGAGTGTGTAATTCGGTTTAAAGAGTCAAGTATTTCCGCT 490
QY 2231 TGAATACCTACGCTCTCTAGTGTATGTGGTGTAGTGATAGACAAAGGGATTCCTGTC 2290
DB 491 TGAATACCTACGCTCTCTAGTGTATGTGGTGTAGTGATAGACAAAGGGATTCCTGTC 550
QY 2291 ACCGAGGCTTAAATTTGAAGCGCGCTTTAAATAT 2325
DB 551 ACCGAGGCTTAAATTTGAAGCGCGCTTTAAATAT 585

RESULT 96
LOCUS DA766261 564 bp mRNA linear EST 09-NOV-2005
DEFINITION DA766261 NTONG2 Homo sapiens cDNA clone NTONG2007406 5', mRNA
sequence.
ACCESSION DA766261
VERSION DA766261.1 GI:81279987
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

QY 481 GGTGAGAACAGAGAAATACATGTTTATCTGAAATTCCTCAATAGCGACGTCGGGACGGTCCGGGC 540
 Db 481 GGTGAGAACAGAGAAATACATGTTTATCTGAAATTCCTCAATAGCGACGTCGGGACGGTCCGGGC 540
 QY 541 GCAGTCTTAATGCTCTCTTGGGAG 564
 Db 541 GCAGTCTTAATGCTCTCTTGGGAG 564

RESULT 97
 DA033029
 LOCUS
 DEFINITION
 DA033029 ASTRO2 Homo sapiens cDNA clone ASTRO2017889 5', mRNA
 sequence.
 DA033029
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saiko, K., Isono, Y., Irie, R., Kishida, N., Yonekawa, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Tanase, T., Nagai, K., Kikuchi, H., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
 Genome Res. 16 (1), 55-65 (2006)

TITLE
 JOURNAL
 PUBMED
 COMMENT

FEATURES
 source
 Location/Qualifiers
 1..563
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="ASTRO2017889"
 /cell_type="normal astrocytes (NHAS732)"
 /note="Vector: pME18SFL3"

ORIGIN
 Query Match 16.4%; Score 511; DB 9; Length 563;
 Best Local Similarity 99.8%; Pred No 1.7e-264;
 Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 AAAGCTCCGAGGCCAGGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGGCGGCTTCGC 67
 Db 2 AAAGCTCCGAGGCCAGGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGGCGGCTTCGC 61
 QY 68 GCGCTGGGTTCTACCGCGCGCGCGCGGAGGAGCCATGCAACAGGACCGGAGTGGG 127
 Db 62 GCGCTGGGTTCTACCGCGCGCGCGCGGAGGAGCCATGCAACAGGACCGGAGTGGG 121
 QY 128 GCGCGCGCAGCATGAAGCGCGCGCGCTCCTCATAGCGACGTCGGGACGGTCCGGGC 187

Db 122 GCGCGCGCAGCATGAAGCGCGCGCGCTCCTATAGCGACGTCGGGACGGTCCGGGC 181
 QY 188 GCGCGCGCGGAGGAGAAATGCAACATGCGAGCGCAATGGAACAGAACAGCTGGGTG 247
 Db 182 GCGCGCGCGGAGGAGAAATGCAACATGCGAGCGCAATGGAACAGAACAGCTGGGTG 241
 QY 248 TTGAGATATTTGAAACTGCGGACTGTTGAGGAGAAATTTGAATCAAGGATCGGCTAAAT 307
 Db 242 TTGAGATATTTGAAACTGCGGACTGTTGAGGAGAAATTTGAATCAAGGATCGGCTAAAT 301
 QY 308 TGGAGCTTTTATGTTGAGCGGTAATCTCGAGTCAAGTCTTAAAGCTGCTTCCCGGATA 367
 Db 302 TGGAGCTTTTATGTTGAGCGGTAATCTCGAGTCAAGTCTTAAAGCTGCTTCCCGGATA 361
 QY 368 CCAGAAATATCATGCTGCTACATGATGCTTAAGGACCAACATGATTTTCATGTTGTGAAGA 427
 Db 362 CCAGAAATATCATGCTGCTACATGATGCTTAAGGACCAACATGATTTTCATGTTGTGAAGA 421
 QY 428 GGNATGATCCAGATCGACCTCAATTCAGACAGAAATCTATTACCTTGCATGCTGTTGAGA 487
 Db 422 GGAATGATCCAGATCGACCTCAATTCAGACAGAAATCTATTACCTTGCATGCTGTTGAGA 481
 QY 488 ACAGAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGGACGAGTCT 547
 Db 482 ACAGAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGGACGAGTCT 541
 QY 548 TAATGCTCTCTTGGAGCGCTCT 569
 Db 542 TAATGCTCTCTTGGAGCGCTCT 563

RESULT 98
 LOCUS
 DEFINITION
 DB020491 TEST12 Homo sapiens cDNA clone TEST12000701 5', mRNA
 sequence.
 DB020491
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saiko, K., Isono, Y., Irie, R., Kishida, N., Yonekawa, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Tanase, T., Nagai, K., Kikuchi, H., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
 Genome Res. 16 (1), 55-65 (2006)

TITLE
 JOURNAL
 PUBMED
 COMMENT

FEATURES
 Location/Qualifiers
 1..559
 /organism="Homo sapiens"
 /mol_type="mRNA"

REFERENCE
AUTHORS
1 (bases 1 to 575)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, I., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, K., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Rakanashi, Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isegai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (11), 55-65 (2006)
16344560
JOURNAL
PUBMED

DA523081 560 bp mRNA linear EST 06-NOV-2005
DA523081 FEBR2 Homo sapiens CDNA clone FEBR2013746 5', mRNA

```
sequence.
ACCESSION   DA523081
VERSION     DA523081.1
KEYWORDS    GI:80930395
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens

REFERENCE   1 (bases 1 to 560)
AUTHORS    Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
            Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
            Genome Res. 16 (1), 55-65 (2006)
            16344560
            Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.

FEATURES             source
    source            1..560
                        Location/Qualifiers
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="FEBR2013746"
                        /tissue_type="brain"
                        /dev_stage="fetal"
                        /clone_lib="FEBR2"
                        /note="Vector: pME18SFL3"

ORIGIN
Query Match      16.3%; Score 509; DB 9; Length 560;
Best Local Similarity 99.8%; Pred. No. 2.1e-263;
Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGCGCTTGC CGCGCTGGGT 76
    |||||
Db 1 GAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGCGCTTGC CGCGCTGGGT 60

Qy 77 TGTCAACCGCGCGCGCGCGGAGGAGCACTGCAACCGAGACCGGAGTGGAGCGCGCA 136
    |||||
Db 61 TGTCAACCGCGCGCGCGCGGAGGAGCACTGCAACCGAGACCGGAGTGGAGCGCGCA 120

Qy 137 GCATGAAGCGCGCGCGCGCGCTCATAGGCGCACGTCGGAGCGGTCGGCGCGCGCGG 196
    |||||
Db 121 GCATGAAGCGCGCGCGCGCGCTCATAGGCGCACGTCGGAGCGGTCGGCGCGCGG 180

Qy 197 GGGAGGAAATGCAACATGCGAGCAATGCAACAGCAAGCAAGCTGGGTGTGAGATAT 256
    |||||
Db 181 GGGAGGAAATGCAACATGCGAGCAATGCAACAGCAAGCAAGCTGGGTGTGAGATAT 240

Qy 257 TTGAACATGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGGCTAAATTTGAGGCTT 316
    |||||
Db 241 TTGAACATGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGGCTAAATTTGAGGCTT 300

Qy 317 TTTATGTGAGCGGTATTCCTGAGTCAAGCTTAAAGAGCTGTCGCGATACCGAATAT 376
    |||||
Db 301 TTTATGTGAGCGGTATTCCTGAGTCAAGCTTAAAGAGCTGTCGCGATACCGAATAT 360

377 ATCATGGCTACATGATGGCTTAGGACACACATGATTTCATGTTTGTGAAGGAGTATC 436
    |||||
361 ATCATGGCTACATGATGGCTTAGGACACACATGATTTCATGTTTGTGAAGGAGTATC 420
    |||||
437 CAGATGAGCTTATTCAGACAGAACTTATACCTTGGCCATCTCTGGTGAGAACAGAGAAA 496
    |||||
421 CAGATGAGCTTATTCAGACAGAACTTATACCTTGGCCATCTCTGGTGAGAACAGAGAAA 480
    |||||
497 ATACACTGTTTATTCAGAAATCCCAAACTATCAATAGACGACGAGTCTTAAATGCTCT 556
    |||||
481 ATACACTGTTTATTCAGAAATCCCAAACTATCAATAGACGACGAGTCTTAAATGCTCT 540
    |||||
557 CTTGGAGCGCTCTTTTGGAT 576
    |||||
541 CTTGGAGCGCTCTTTTGGAT 560
    |||||

RESULT 101
LOCUS       DA437193
DEFINITION DA437193 CTONG2 Homo sapiens cDNA clone CTONG2004408 5', mRNA
ACCESSION   DA437193
VERSION     DA437193.1
KEYWORDS    GI:80891371
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
            1 (bases 1 to 611)
            Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
            Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
            Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H.,
            Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
            Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
            Diversification of Transcriptional Modulation: Large-scale
            Identification and Characterization of Putative Alternative
            Promoters of Human Genes
            Genome Res. 16 (1), 55-65 (2006)
            16344560
            Contact: Takao Isogai
            FLJ Project (HRI Team)
            Helix Research Institute
            2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: flj-cdna@nifty.com
            NEDO human cDNA project (New Energy and Industrial Technology
            Developmental Organization, Japan); cDNA library construction:
            Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
            Research Association for Biotechnology (RAB) and Biotechnology
            Center, National Institute of Technology and Evaluation; 3'-end one
            pass sequencing: RAB.

FEATURES             source
    source            1..611
                        Location/Qualifiers
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="CTONG2004408"
                        /tissue_type="tongue"
                        /clone_lib="CTONG2"
                        /note="Vector: pME18SFL3"

ORIGIN
Query Match      16.3%; Score 509; DB 9; Length 611;
Best Local Similarity 99.7%; Pred. No. 2.1e-263;
Matches 609; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 54 AGTCCCGCTTCCCGCGCTGGTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 113
    |||||
```

Db 1 AGTCCCGCTTCGCGCGCTGGGTTGTCACCGCGCGCGCGCGAGGAGCCACTGCAACC 60

Qy 114 AGGACCGGAGTGGAGCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATGAGCCAGCTC 173

Db 61 AGGACCGGAGTGGAGCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATGAGCCAGCTC 120

Qy 174 GGGACCGTCCGCGCGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 233

Db 121 GGGACCGTCCGCGCGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180

Qy 234 AGACAGCTGGGCTGTTCAGATATTTGAAACTGCGGACTGTGAGGAGGAGGAGGAGGAGG 293

Db 181 AGACAGCTGGGCTGTTCAGATATTTGAAACTGCGGACTGTGAGGAGGAGGAGGAGGAGG 240

Qy 294 GGATCGGCTAAATTTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAA 353

Db 241 GGATCGGCTAAATTTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAA 300

Qy 354 GCTGCTTCGGCATACAGAAATATCATGCTACATGATGGCTAAGGCAACCATGATTT 413

Db 301 GCTGCTTCGGCATACAGAAATATCATGCTACATGATGGCTAAGGCAACCATGATTT 360

Qy 414 CATGTTTGTGAAGAGGAGTATCCAGATGGACCTCATTCAGACAGGATCTTATACCTGTC 473

Db 361 CATGTTTGTGAAGAGGAGTATCCAGATGGACCTCATTCAGACAGGATCTTATACCTGTC 420

Qy 474 CATGCTGTGAGAGCAGAGAGAAATACACTCTGTTTATTCGAAATCCCAAACTATCAA 533

Db 421 CATGCTGTGAGAGCAGAGAGAAATACACTCTGTTTATTCGAAATCCCAAACTATCAA 480

Qy 534 TAGACAGCAGCTCTTAATCTCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCAACT 593

Db 481 TAGACAGCAGCTCTTAATCTCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCAACT 540

Qy 594 GGCATATGGAATGATTTTCGAGAGAGAACTATTAAGAGAGAGAGAGAGAGAGAGAGAG 653

Db 541 GGCATATGGAATGATTTTCGAGAGAGAACTATTAAGAGAGAGAGAGAGAGAGAGAGAG 600

Qy 654 AGTCGGAATTG 664

Db 601 AGTCGGAATTG 611

RESULT 102

DB041170

LOCUS

DEFINITION

DB041170 TESTI2 Homo sapiens cDNA clone BRTH1000306 5', mRNA

ACCESSION

DB041170

VERSION

DB041170.1 GI:81223330

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 547)

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi, F., Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.

TITLE

Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes

JOURNAL

PUBMED

16344560

CONTACT: Takao Isogai

FLJ Project (HRI Team)

Helix Research Institute

2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: flj-cdn@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: HRI, Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

Location/Qualifiers

1. 547

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="TESTI2028180"

/tissue_type="testis"

/clone_id="TESTI2"

/note="Vector: pME18SFL3"

FEATURES

source

ORIGIN

Query Match 16.3%; Score 508; DB 9; Length 547;

Best Local Similarity 100.0%; Pred. No. 7.2e-263;

Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1332 TGCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAAC 1391

Db 40 TGCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAAC 99

Qy 1392 TGAATTTATTTATCCCAAGTAGAGATGATGTTATGGAAGGCGAGACTCATTTGATGAGT 1451

Db 100 TGAATTTATTTATCCCAAGTAGAGATGATGTTATGGAAGGCGAGACTCATTTGATGAGT 159

Qy 1452 GCCTGATTTCTGACGCCACTAATTTATCTATGAAGAAACACACACATCTCGATAAATAT 1511

Db 160 GCCTGATTTCTGACGCCACTAATTTATCTATGAAGAAACACACACATCTCGATAAATAT 219

Qy 1512 CCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAGGAAATTTGATTTTATTTTGC 1571

Db 220 CCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAGGAAATTTGATTTTATTTTGC 279

Qy 1572 CTCTGAATGCAAAACACAGGTTTCCTGTCATTTATACAAAATACATCTATTTTAAAGGAAG 1631

Db 280 CTCTGAATGCAAAACACAGGTTTCCTGTCATTTATACAAAATACATCTATTTTAAAGGAAG 339

Qy 1632 CAATATAAAGCATCCAGTGGTGGCTGCTCCCAAGTGAATTTCAAGTGTCTTATCAAA 1691

Db 340 CAATATAAAGCATCCAGTGGTGGCTGCTCCCAAGTGAATTTCAAGTGTCTTATCAAA 399

Qy 1692 AGAGGAGATAGCAATTAACCATGTTGTAATGGGAAGTTCTTGGCGGCATGATCTAATAT 1751

Db 400 AGAGGAGATAGCAATTAACCATGTTGTAATGGGAAGTTCTTGGCGGCATGATCTAATAT 459

Qy 1752 CCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCGCAAAAGACTCCCTTTAGA 1811

Db 460 CCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCGCAAAAGACTCCCTTTAGA 519

Qy 1812 GCATCACTGTACGTAGTCAGTTACGTA 1839

Db 520 GCATCACTGTACGTAGTCAGTTACGTA 547

RESULT 103

DB041170

LOCUS

DEFINITION

DB041170 TESTI2 Homo sapiens cDNA clone BRTH1000306 5', mRNA

ACCESSION

DB041170

VERSION

DB041170.1 GI:80866355

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE
AUTHORS
1 (bases 1 to 559)
Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
FEATURES
source
Location/Qualifiers
1..559
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRHA1000306"
/tissue_type="thalamus"
/clone_lib="BRTHAL"
/notes="Vector: pME18SFL3"
ORIGIN
Query Match 16.3%; Score 508; DB 9; Length 559;
Best Local Similarity 99.8%; Pred. No. 7.2e-263;
Matches 558; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 8 AAGCTTCGAGCCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTCGCGCTCGC 67
DB 1 AAGCTTCGAGCCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTCGCGCTCGC 60
QY 68 CGCTTGGGTGTGACCGCGCGCGCGCGAGGAGCCACTGCAACGACCGAGTGA 127
DB 61 CGCTTGGGTGTGACCGCGCGCGCGCGAGGAGCCACTGCAACGACCGAGTGA 120
QY 128 GCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGACGGTCCGGGC 187
DB 121 GCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGACGGTCCGGGC 180
QY 188 GGGGCGGGGGAAGGAAATGCAACATGCGCAGCGCAATGGAAACAGAACAGCTGGGTG 247
DB 181 GGGGCGGGGGAAGGAAATGCAACATGCGCAGCGCAATGGAAACAGAACAGCTGGGTG 240
QY 248 TTGAGATATTTGAACTGCGGCTGTCGAGGAGATTTGATTCACAGGATCGGCTAAAT 307
DB 241 TTGAGATATTTGAACTGCGGCTGTCGAGGAGATTTGATTCACAGGATCGGCTAAAT 300
QY 308 TGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGTTAAAAAGCTGCTTCCGATA 367
DB 301 TGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGTTAAAAAGCTGCTTCCGATA 360
QY 368 CCAGAAATATCATGCGCTACATGCTTAAGCCACCAATGATTTCTATGTTTGAAGA 427
DB 361 CCAGAAATATCATGCGCTACATGCTTAAGCCACCAATGATTTCTATGTTTGAAGA 420
QY 428 GGAATGATCCAGATGGACCTCATTCAGACAGATCTTATTCCTGCCATGCTGCTGGAGA 487
DB 421 GGAATGATCCAGATGGACCTCATTCAGACAGATCTTATTCCTGCCATGCTGCTGGAGA 480
QY 488 ACAGAGAAATACACTGTTTATTCGAAATTTCCCAAAACTCATCAATAGCAGCAGTCT 547

Db 481 ACAGAGAAATACACTGTTTATTCGAAATTTCCCAAAACTCATCAATAGCAGCAGTCT 540
QY 548 TAATGCTCTCTTGGAGCC 566
DB 541 TAATGCTCTCTTGGAGCC 559

RESULT 104
DA956211
LOCUS DA956211 SPLEN2 Homo sapiens cDNA clone SPLEN2035091 5', mRNA
DEFINITION DA956211 559 bp mRNA linear EST 09-NOV-2005
sequence.
ACCESSION DA956211
VERSION DA956211.1 GI:81289034
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 559)
Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
FEATURES
source
Location/Qualifiers
1..559
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRHA1000306"
/tissue_type="thalamus"
/clone_lib="BRTHAL"
/notes="Vector: pME18SFL3"
ORIGIN
Query Match 16.3%; Score 508; DB 9; Length 559;
Best Local Similarity 99.8%; Pred. No. 7.2e-263;
Matches 558; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1454 CTGATTCGTGTCACGCCCTAATTCATATGAAAGAAACAACAGACATCTCGATAATCC 1513
DB 1 CTGATTCGTGTCACGCCCTAATTCATATGAAAGAAACAACAGACATCTCGATAATCC 60
QY 1514 ATGCATCTTTCATGTTTTTCCCAAGTCAAGAGGAAATTCAGTTTATTTTGGCT 1573
DB 61 ACGATCTTTCATGTTTTTCCCAAGTCAAGAGGAAATTCAGTTTATTTTGGCT 120
QY 1574 CTGATTCGAAACAGGTTTCCGTCATTTATACAAATATCATCTATTTAAAGGAAAGCA 1633
DB 121 CTGATTCGAAACAGGTTTCCGTCATTTATACAAATATCATCTATTTAAAGGAAAGCA 180

```



```
QY 1634 ATATATAACGATCCAGTGGTGGCTGCTCCAGTGNATTTCAAGTGTCTCTATCAAG 1693
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1694 AGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCC 1753
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1754 AAGTGTAGTGAAGTCCAGAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGC 1813
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 AAGTTTGAAGTCCAGAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGC 360
QY 1814 ATCACCTGTACGTAGTCAAGTATACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGGTG 1873
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1874 GCTACTCATCTTGTGTGCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTA 1933
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1934 ACCAGAAGATCCACACTGTGTGTCCTTTTACAAGTATCAAGTCTCTGAAGATGACCCAA 1993
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1994 CTTGCAAAACAAAGGAATT 2012
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 CTTGCAAAACAAAGGAATT 559
```

RESULT 105

```
DA450229 DA450229 CTONG2 Homo sapiens cDNA clone CTONG2022308 5', mRNA EST 06-NOV-2005
LOCUS sequence.
DEFINITION DA450229
ACCESSION DA450229
VERSION DA450229.1 GI:80908670
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 553)
```

TITLE

Promoters of Human Genes

Genome Res. 16 (1), 55-65 (2006)

JOURNAL

16344560

COMMENT

Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation
pass sequencing: RAB.

FEATURES

source

1...553
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2022308"

/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 16.1%; Score 502; DB 9; Length 553;
Best Local Similarity 99.8%; Pred. No. 1.3e-259;
Matches 552; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 AAGTGTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCCCGCCGCTGCTTTTAAAGTCCG 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1 AAGTGTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCCCGCCGCTGCTTTTAAAGTCCG 60
QY 61 CGTTCCGCCCTTGGGTTGTCAACCGCGCGCGCGCGGAGGAGCCACTGCAACACGAGACCG 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 CGTTCCGCCCTTGGGTTGTCAACCGCGCGCGCGGAGGAGCCACTGCAACACGAGACCG 120
QY 121 GAGTGAGCGCGCGCAGCATGAAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGAGCGG 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 GAGTGAGCGCGCGCAGCATGAAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGAGCGG 180
QY 181 TCCGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAACACGAGACAG 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 TCCGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAACACGAGACAG 240
QY 241 CTGGGTGTTGAGATATTTGAAACTCGCGACTGTGAGGAGATAATTGAATCACAGGATCGG 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 CCTAAATTCGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTT 360
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 CCTAAATTCGAGCCTTTTATGTTGAGGTGATTCCTGGAGTCAGCTTAAAGCTGCTT 360
QY 361 GCCGATACCAGAAAATATCATGGCTACATGATGGCTAAGGCACCAATGATTTTCATGTTT 420
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 421 GTGAGAGGAATGATCCAGATGGACCTCATTCACAGCAATCTATTACCTTGCCATGCT 480
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 481 GTGAGAAACAGAGAAAATACACTGTTTTTATCTGAAATTCCTCAAACTATCAATAGACGA 540
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 481 GGTGAGAACAGAGAAAATACACTGTTTTTATCTGAAATTCCTCAAACTATCAATAGACGA 540
QY 541 GCAGTCTTAATGC 553
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 541 GCAGTCTTAATGC 553
```

RESULT 106

DA831627 DA831627 PLACE1 Homo sapiens cDNA clone PLACE1007416 5', mRNA
LOCUS sequence.
DEFINITION DA831627
ACCESSION DA831627
VERSION DA831627.1 GI:82358835
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

AUTHORS

1 (bases 1 to 631)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes

TITLE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 631)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes

Db	183	AAATGCAACATGGCGAGCAGCAATGGAACAGACAGCTGGGTCTTCAGATATTTTGAACCT	242
Qy	265	GGGACTGTGGAGAGAAATTAATCAAGATCGGCTTAATTTGGAGCCTTTTATGTT	324
Db	243	GGGACTGTGGAGAGAAATTAATCAAGATCGGCTTAATTTGGAGCCTTTTATGTT	302
Qy	325	GAGCGTATTCCTGAGTCAAGTAAAGCTGCTCCGATACACAGAAATATCATGGC	384
Db	303	GAGCGTATTCCTGAGTCAAGTAAAGCTGCTCCGATACACAGAAATATCATGGC	362
Qy	385	TACATGATGCTTAAGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGA	444
Db	363	TACATGATGCTTAAGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGA	422
Qy	445	CCTCATTCAGACAGAAATCTATTACCTTCCATCTGCTGAGAGAAATATACACTG	504
Db	423	CCTCATTCAGACAGAAATCTATTACCTTCCATCTGCTGAGAGAAATATACACTG	482
Qy	505	TTTTTATTCGAAATTC	520
Db	483	TTTTTATTCGAAATTC	498
RESULT 109			
LOCUS	AA496257	500 bp	mRNA
DEFINITION	aa24b08.s1 NCI CGAP GCBI Homo sapiens cDNA clone IMAGE:814167 3'		linear EST 18-AUG-1997
ACCESSION	AA496257		similar to SW:DP66 HUMAN P42658 DIPEPTIDYL PEPTIDASE IV LIKE
VERSION	AA496257.1	GI:2229578	PROTEIN i, mRNA sequence.
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS			Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
TITLE			Hominidae; Homo.
JOURNAL			1 (bases 1 to 500)
COMMENT			NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D. CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
FEATURES			Possible reversed clone: similarity on wrong strand
source			Seq primer: -40m13 fwd. ET from Amersham
			High quality sequence stop: 405.
FEATURES			Location/Qualifiers
source			1..500
			/organism="Homo sapiens"
			/mol_type="mRNA"
			/db_xref="GDB:6032116"
			/db_xref="taxon:9606"
			/clone="IMAGE:814167"
			/tissue_type="germinal center B cell"
			/lab_host="DH10B"
			/clone_lib="NCI CGAP GCBI"
			/notes="Vector: pT7T3D-Paci; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
Qy	368	CCAGAAAATATCATGGCTACATGATGGCTAAGGCACACATGATTTTCATGTTTGAAGA	427
Db	373	CCAGAAAATATCATGGCTACATGATGGCTAAGGCACACATGATTTTCATGTTTGAAGA	432
Qy	428	GGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCTGGTGAGA	487
Db	433	GGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCTGGTGAGA	492
Qy	488	ACAGAGAAAATACACTGTTT	507
Db	493	ACAGAGAAAATACACTGTTT	512
RESULT 108			
LOCUS	CN427226	704 bp	mRNA
DEFINITION	1700045364785 GRN_ES Homo sapiens cDNA 5', mRNA sequence.		linear EST 16-MAY-2004
ACCESSION	CN427226		
VERSION	CN427226.1	GI:47414820	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS			Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
TITLE			Hominidae; Homo.
JOURNAL			1 (bases 1 to 704)
PUBMED			Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
COMMENT			Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation Nat. Biotechnol. 22 (6), 707-716 (2004) 15146197 Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@geron.com Insert Length: 704 Std Error: 0.00. Location/Qualifiers 1..704 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /tissue_type="embryonic stem cells, cell lines H1, H7, and H9" /clone_lib="GRN_ES" /note="Oligo dT primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"
FEATURES			
source			
ORIGIN			
Query Match		15.9%;	Score 496; DB 8; Length 704;
Best Local Similarity		100.0%;	Pred. No. 2.4e-256; Indels 0; Gaps 0;
Matches		496; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	25	GGCGGTGTACTGCGCGCTGCTTCTTAGTGCCTGCTTCCGCGCTGGTGTGTCACCG	84
Db	3	GGCGGTGTACTGCGCGCTGCTTCTTAGTGCCTGCTTCCGCGCTGGTGTGTCACCG	62
Qy	85	GGCGGTGTACTGCGCGCTGCTTCTTAGTGCCTGCTTCCGCGCTGGTGTGTCACCG	144
Db	63	GGCGGTGTACTGCGCGCTGCTTCTTAGTGCCTGCTTCCGCGCTGGTGTGTCACCG	122
Qy	145	GGCGGTGTACTGCGCGCTGCTTCTTAGTGCCTGCTTCCGCGCTGGTGTGTCACCG	204
Db	123	GGCGGTGTACTGCGCGCTGCTTCTTAGTGCCTGCTTCCGCGCTGGTGTGTCACCG	182
Qy	205	AAATGCAACATGGCGAGCAGCAATGGAACAGACAGCTGGGTCTTCAGATATTTTGAACCT	264

[5'-TGTACCAATCTGAAGTGGAGCGCGCTCATTTTCTTTTCTTTT-3',
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Donaldso."

ORIGIN
Query Match 15.9%; Score 495; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 81e-256;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2625 TGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAAATGTCATTTTGACCA 2684
Db 495 TGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAAATGTCATTTTGACCA 436
QY 2685 TACCAGTATATTACTGAGTTTCTTGTAGTGGGCTGGAAAGCCATATGATTACAGATCTA 2744
Db 435 TACCAGTATATTACTGAGTTTCTTGTAGTGGGCTGGAAAGCCATATGATTACAGATCTA 376
QY 2745 TCCTCAGGAGACACAGCAATGAGTTCTGTAATCGGGAACATTTATGAATCTCATCT 2804
Db 375 TCCTCAGGAGACACAGCAATGAGTTCTGTAATCGGGAACATTTATGAATCTCATCT 316
QY 2805 TTGCACTACCTTCAAGAAACCTTGGATCACGTATTGCTGCTCTAAAGTGAATAAT 2864
Db 315 TTGCACTACCTTCAAGAAACCTTGGATCACGTATTGCTGCTCTAAAGTGAATAAT 256
QY 2865 TTGACCTGTGTAGAACTCTCTGSPATACACTGCTATTAAACCAATGAGAGGTTTAAT 2924
Db 255 TTGACCTGTGTAGAACTCTCTGSPATACACTGCTATTAAACCAATGAGAGGTTTAAT 196
QY 2925 CAACAGAAACACAGAAATGATCATCATTTTGTATGATCTGATGATGATGATGAT 2984
Db 195 CAACAGAAACACAGAAATGATCATCATTTTGTATGATCTGATGATGATGATGAT 136
QY 2985 GAAATAAATGTGTGTCATGAGGGGTCTACGGTTTGTGAGTAACTATACCTTAA 3044
Db 135 GAAATAAATGTGTGTCATGAGGGGTCTACGGTTTGTGAGTAACTATACCTTAA 76
QY 3045 CCCACATGCTCAAAATCAATGATACATATTTCTGAGAGCCAGCAATACCAATAGAA 3104
Db 75 CCCACATGCTCAAAATCAATGATACATATTTCTGAGAGCCAGCAATACCAATAGAA 16
QY 3105 TTACTAAAAA 3119
Db 15 TTACTAAAAA 1

RESULT 110
BE789604
LOCUS 601481552F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3884197 5',
DEFINITION mRNA sequence.
ACCESSION BE789604
VERSION BE789604.1 GI:10210802
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo;
1 (bases 1 to 617)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/BTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Cloned by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9657 row: 9 column: 14
High quality sequence stop: 614.
Location/Qualifiers
1. 617
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3884197"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_68"
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 15.8%; Score 493; DB 7; Length 617;
Best Local Similarity 99.8%; Pred. No. 9.9e-255;
Matches 613; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2213 GAGTCAAGTATTTTCGCTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTCATAG 2272
Db 1 GAGTCAAGTATTTTCGCTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTCATAG 60
QY 2273 ACAACAGGGATTCCTGTACCGAGGGCTTAAATTTGAAGCGCTTTTAAATATAAAATGG 2332
Db 61 ACAACAGGGATTCCTGTACCGAGGGCTTAAATTTGAAGCGCTTTTAAATATAAAATGG 120
QY 2333 GTCAA-ATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGAT 2391
Db 121 GTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGAT 180
QY 2392 TTCAATTGACTTAGATTCGTTGGGCATCCAGCGCTGGTCTCTATGAGGATACCTTCCCTG 2451
Db 181 TTCAATTGACTTAGATTCGTTGGGCATCCAGCGCTGGTCTCTATGAGGATACCTTCCCTG 240
QY 2452 ATGCGATTAATGACAGAGTCTAGATATCTTTCAGGTTTGTCTTGTGGGGCCAGCTCACT 2511
Db 241 ATGCGATTAATGACAGAGTCTAGATATCTTTCAGGTTTGTCTTGTGGGGCCAGCTCACT 300
QY 2512 CTGTGGATCTTCTATGATACAGGATACAGGAACTTATGCGTACCGCTGACCGCAAT 2571
Db 301 CTGTGGATCTTCTATGATACAGGATACAGGAACTTATGCGTACCGCTGACCGCAAT 360
QY 2572 GAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAGTTCCCTCTGAAACCA 2631
Db 361 GAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAGTTCCCTCTGAAACCA 420
QY 2632 AATCGTTTACTGCTCTTACATGCTTCTGATGAGAAATGTCATTTTGCACATACCAGT 2691
Db 421 AATCGTTTACTGCTCTTACATGCTTCTGATGAGAAATGTCATTTTGCACATACCAGT 480
QY 2692 ATATTACTGAGTTTCTTAGTGGGCTGGAAGCCATATGATTTTACAGATCTATCCTCAG 2751
Db 481 ATATTACTGAGTTTCTTAGTGGGCTGGAAGCCATATGATTTTACAGATCTATCCTCAG 540
QY 2752 GAGAGACACAGCATTAAGAGTTCTGAAATCGGAGAACATTTATGAATCTGTCATCTTTTGCAC 2811
Db 541 GAGAGACACAGCATTAAGAGTTCTGAAATCGGAGAACATTTATGAATCTGTCATCTTTTGCAC 600
QY 2812 TACCTTCAAGAAA 2825
Db 601 TACCTTCAAGAAA 614

RESULT 111
DA720650
LOCUS DA720650 NT2R13 Homo sapiens cDNA clone NT2R13000840 5', mRNA
DEFINITION sequence.

```

ACCESSION DA720650
VERSION DA720650.1 GI:82366371
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 492)
AUTHORS Kimura.K., Wakamatsu.A., Suzuki.Y., Ota.T., Nishikawa.T.,
Yamashita.R., Yamamoto.J., Sekine.M., Tsuritani.K., Wakaquri.H.,
Ishii.S., Sugiura.T., Saito.K., Isono.Y., Irie.R., Kuehida.N.,
Yoneyama.T., Otsuka.R., Kanda.K., Yokoi.T., Kondo.H., Wagatsuma.M.,
Murakawa.K., Ishida.S., Iehibashi.T., Takahashi-Fujii.A.,
Tanase.T., Nagai.K., Kikuchi.H., Nakai.K., Isogai.T. and Sugano.S.
TITLE Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
FEATURES
source
location/Qualifiers
1..492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2R13000840"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2R13"
/notes="Vector: pME18SFL3; majorly NT2 neuron; mRNA from
NT2 neuronal precursor cells treated 2-weeks mitotic
inhibitor after 5-weeks retinoic acid (RA) induction"
ORIGIN
Query Match 15.8%; Score 492; DB 9; Length 492;
Best Local Similarity 100.0%; Pred. No. 3.4e-254;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGTGCTAAAGCTCCGAGGCCGAGCGCGTCTACTGCGCGCGCTCTTTAGTGGCG 60
DB 1 AAGTGCTAAAGCTCCGAGGCCGAGCGCGTCTACTGCGCGCGCTCTTTAGTGGCG 60
QY 61 CGTTGCCGCTTGGGTTGTCTACCGCGCGCGCGCGCGCGCTCTTCAACAGACCG 120
DB 61 CGTTGCCGCTTGGGTTGTCTACCGCGCGCGCGCGCGCGCTCTTCAACAGACCG 120
QY 121 GAGTGGAGCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCCTCGGACCG 180
DB 121 GAGTGGAGCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCCTCGGACCG 180
QY 181 TCGGCGCGCGCGCGGGAAGAAATGCAACATGGCAGCAGCAATGGAAAACAGACAG 240
DB 181 TCGGCGCGCGCGCGGGAAGAAATGCAACATGGCAGCAGCAATGGAAAACAGACAG 240
QY 241 CTGGGTGTTGAGTATTTGAACTCGGACTGTGAGGAGATATTTGAATCACAGGATCG 300
DB 241 CTGGGTGTTGAGTATTTGAACTCGGACTGTGAGGAGATATTTGAATCACAGGATCG 300
QY 301 CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGCTTAAAGAGCTGCTT 360

```

```

Db 301 CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGCTTAAAGAGCTGCTT 360
QY 361 GCCGATACAGAAAATATCATGGCTACATGATGGCTAAGGACGACCATGATTTTCATGTTT 420
Db 361 GCCGATACAGAAAATATCATGGCTACATGATGGCTAAGGACGACCATGATTTTCATGTTT 420
QY 421 GTGAAGAGGATGATCATCAGATGACCTATTTCAGACAAATCTATTACCTTGGCCATGCT 480
Db 421 GTGAAGAGGATGATCATCAGATGACCTATTTCAGACAAATCTATTACCTTGGCCATGCT 480
QY 481 GGTGAGACACAGA 492
Db 481 GGTGAGACACAGA 492
RESULT 112
LOCUS BG165205 602344074F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4453987 5',
mRNA sequence.
VERSION BG165205 GI:12671908
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 639)
AUTHORS NIH-MGC Hsttp://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10245 row: d column: 20
High quality sequence stop: 639.
FEATURES
source
location/Qualifiers
1..639
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4453987"
/tissue_type="hypernephroma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 89"
/notes="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
ORIGIN
Query Match 15.5%; Score 485; DB 2; Length 639;
Best Local Similarity 99.8%; Pred. No. 2.2e-250;
Matches 605; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1878 CTCACATCTTGTGTCATCAGTCAGTCAGTCCTTCTTATTAAGTAAGTATAGTAACCA 1937
Db 1 CTCACATCTTGTGTCATCAGTCAGTCAGTCCTTCTTATTAAGTAAGTATAGTAACCA 60
QY 1938 GAAGAATCCACATCTGTGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTG 1997
Db 61 GAAGAATCCACATCTGTGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTG 120
QY 1998 CAAACAAAGAAATTTGGGCGCACCATTTTGGATTGAGATTCAGCAGGTCTCTCTCGACTATAC 2057

```

Db 121 CAACACAAAGGAATTTTGGGCCACCA-TTGGATTACGACGGTCTCTTCTGACTATAC 179
QY 2058 TCCTCCGAAATTTCTCTTTTGAAGTACTACTGGATTTACATTGATGGATGCTCTA 2117
Db 180 TCCTCCGAAATTTCTCTTTTGAAGTACTACTGGATTTACATTGATGGATGCTCTA 239
QY 2118 CAAGCCTCATGATCTACAGCCTCGAAGAAATATCCTACTGTGCTGTTTATATATGTTG 2177
Db 240 CAAGCCTCATGATCTACAGCCTCGAAGAAATATCCTACTGTGCTGTTTATATATGTTG 299
QY 2178 TCCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTGAATAC 2237
Db 300 TCCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTGAATAC 359
QY 2238 CTTAGCCTCTTAGGTTATGTTGTTAGTAGTACACAGCGGATCTGTCCACGAGG 2297
Db 360 CTTAGCCTCTTAGGTTATGTTGTTAGTAGTACACAGCGGATCTGTCCACGAGG 419
QY 2298 GCTTAAATTTGAAGCGCCTTTAAATATAAATGAGTCAATAGAAATTTGACGATCAGGT 2357
Db 420 GCTTAAATTTGAAGCGCCTTTAAATATAAATGAGTCAATAGAAATTTGACGATCAGGT 479
QY 2358 GGAAGGACTCCAATATCTAGCTTCTCGATATGTTTATGATGATGATGATGATGATGATGAT 2417
Db 480 GGAAGGACTCCAATATCTAGCTTCTCGATATGTTTATGATGATGATGATGATGATGATGAT 539
QY 2418 CCAGCGCTGTCCTATGAGGATGATCTCTCTGATGAGGATTAATGCGAGGTCAGATAT 2477
Db 540 CCAGCGCTGTCCTATGAGGATGATCTCTCTGATGAGGATTAATGCGAGGTCAGATAT 599
QY 2478 CTTACG 2483
Db 600 CTTACG 605

RESULT 113

LOCUS DA522231
DEFINITION DA522231 FEBRA2 Homo sapiens cDNA clone FEBRA2012626 5', mRNA
VERSION DA522231
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Teuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi,Fuji,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
PLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEBO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); S-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.

FEATURES

source
1..553
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FEBRA2012626"
/tissue_type="brain"
/dev_stage="fetal"
/clone_lib="FEBRA2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 15 4%; Score 481; DB 9; Length 553;
Best Local Similarity 99.8%; Pred. No. 3.2e-248;
Matches 531; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 54 AGTGGCGGTTTCGGCGGCTGGTTGTCTACCGGCGCCCGCCGAGGAGCACTGCAACC 113
Db 1 AGTGGCGGTTTCGGCGGCTGGTTGTCTACCGGCGCCCGCCGAGGAGCACTGCAACC 60
QY 114 AGGACCGGAGTGGAGCGCGCCGACGATGAAGCGCGCGCCGCTCCATAGCGCACGTC 173
Db 61 AGGACCGGAGTGGAGCGCGCCGACGATGAAGCGCGCGCGCCGCTCCATAGCGCACGTC 120
QY 174 GGGACGGTCCGGCGCGCGCGCGGGGGAAGGAAATGCAACATGGCAGCAATGGAAC 233
Db 121 GGGACGGTCCGGCGCGCGCGCGGGGGAAGGAAATGCAACATGGCAGCAATGGAAC 180
QY 234 AGACAGCTGGGTTTGAGATATTTGAACTCGGAGCTGTGAGGAGAAATATTGAATCACA 293
Db 181 AGACAGCTGGGTTTGAGATATTTGAACTCGGAGCTGTGAGGAGAAATATTGAATCACA 240
QY 294 GGATCGGCTAAATTTGAGCGCTTTTATGTGAGCGGTATTCTTGAGTACGCTTAAAAA 353
Db 241 GGATCGGCTAAATTTGAGCGCTTTTATGTGAGCGGTATTCTTGAGTACGCTTAAAAA 300
QY 354 GCTGCTTCCGATACACAGAAAAATATCATGGCTACATGATGGCTAAGGCACCAATGTT 413
Db 301 GCTGCTTCCGATACACAGAAAAATATCATGGCTACATGATGGCTAAGGCACCAATGTT 360
QY 414 CATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTGC 473
Db 361 CATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTGC 420
QY 474 CATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTGC 533
Db 421 CATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTGC 480
QY 534 TAGAGCAGCAGTCTTAAATGCTCTCTTGAAGCGCTTTTGGATCTTTTTCAG 585
Db 481 TAGAGCAGCAGTCTTAAATGCTCTCTTGAAGCGCTTTTGGATCTTTTTCAG 532

RESULT 114

LOCUS BI710784
DEFINITION BI710784 536 bp mRNA linear EST 11-MAR-2002
BI710784 5' similar to TR.075273 OY5273 R26984_1; mRNA sequence.
ACCESSION BI710784
VERSION BI710784.1 GI:15686479
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo
1 (bases 1 to 536)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,A., Marra,M., Pape,D., Wyllie,T., Martin,J., Blisfain,A., Schmitt,A., Theising,B., Ritter,S., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Teagareishvili,R., Williams,I., Jackson,Y. and Bowers,Y.

Thu Jun 22 09:04:26 2006

```

TITLE      Endocrine Pancreas Consortium
JOURNAL    Unpublished (2000)
COMMENT    Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
          Endocrine Pancreas Consortium
          Harvard University, Howard Hughes Medical Institute
          Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
          MA 02138
          Tel: 617-495-1812
          Fax: 617-495-8557
          Email: dmelton@biohp.harvard.edu
          Library was constructed by Dr. J. Ferrer In vivo mass-excised to
          pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
          University Genome Sequencing Center For information on obtaining a
          clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
          Seq primer: -40RP from Gibco
          High quality sequence stop: 428.

FEATURES   Location/Qualifiers
            1..536
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5023286"
             /tissue_type="insulinoma"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="Human insulinoma"
             /note="Organ: pancreas; Vector: pBluescript SK-; Site 1:
             XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system
             (stratagene) by Dr. J. Ferrer, in vivo mass-excised to
             pBluescript SK- by Dr. H. Inoue following the Washington
             University protocol
             (http://genome.wustl.edu/est/lambda_protocol.shtml).
             Please contact Hiroshi Inoue, MD/PhD for further
             information on this library (Metabolism Division, Permutt
             Laboratory, Washington University School of Medicine, Box
             8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
             is a Washington University Pancreas EST project library."

ORIGIN
Query Match      15.4%; Score 480; DB 2; Length 536;
Best Local Similarity 99.8%; Pred. No. 1.1e-247;
Matches 530; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2532 AGGATACAGGACGTTATATGGTCAACCTGACAGAAATGAACAGGCTATTACTTAGG 2591
DB 6 AGGATACAGGACGTTATATGGTCAACCTGACAGAAATGAACAGGCTATTACTTAGG 65
QY 2592 ATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAATGTTTACTGCTTTACA 2651
DB 66 ATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAATGTTTACTGCTTTACA 125
QY 2652 TGGTTTCTGGATGAGAAATGTCATTTTGCATACCATGATATTTACTGAGTTTGTAGT 2711
DB 126 TGGTTTCTGGATGAGAAATGTCATTTTGCATACCATGATATTTACTGAGTTTGTAGT 185
QY 2712 GAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGACACAGCAATGAGT 2771
DB 186 GAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGACACAGCAATGAGT 245
QY 2772 TCCTGAATCGGGAGAACATTATGAACTGTCATCTTTTGCACTACCTTCAAGAAACCTTGG 2831
DB 246 TCCTGAATCGGGAGAACATTATGAACTGTCATCTTTTGCACTACCTTCAAGAAACCTTGG 305
QY 2832 ATCAGGATGTCGTCCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATA 2891
DB 306 ATCAGGATGTCGTCCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATA 365
QY 2892 CACTGGCTATTATTAACAAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATCA 2951
DB 366 CACTGGCTATTATTAACAAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATCA 425
QY 2952 CATTTTGATACCTGGCAGTAACTACTCTCTGAAATAAATGTTGGTGCATGCGAGGGG 3011
DB 426 CATTTTGATACCTGGCAGTAACTACTCTCTGAAATAAATGTTGGTGCATGCGAGGGG 485

```

```

QY 3012 TCTACGGTTTGGTAGTAATCTAATACCTTAACCCACATGCTCAAATC 3062
DB 486 TCTACGGTTTGGTAGTAATCTAATACCTTAACCCACATGCTCAAATC 536

RESULT 115
BI084885/c 855 bp mRNA linear EST 20-JUN-2001
LOCUS 602869453T1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013996 3',
DEFINITION mRNA sequence.
ACCESSION BI084885
VERSION BI084885.1 GI:14503215
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Hominidae; Homo.
REFERENCE 1 (bases 1 to 855)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: CLONTECH Laboratories, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLC1820 row: j column: 13
          High quality sequence start: 3
          High quality sequence stop: 844.
          Location/Qualifiers
            1..855
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5013996"
             /tissue_type="epidermoid carcinoma, cell line"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC_102"
             /note="Organ: salivary gland; Vector: pOTB7; Site 1: XhoI;
             Site 2: EcoRI; cDNA made by oligo-dT priming
             Directionally cloned into EcoRI/XhoI sites using the
             following 5' adaptor: GGCACGAG(G). Library constructed
             by Ling Hong in the laboratory of Gerald M. Rubin
             (University of California, Berkeley) using ZAP-cDNA
             synthesis kit (Stratagene) and Superscript II RT (Life
             Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Query Match      15.4%; Score 479; DB 2; Length 855;
Best Local Similarity 100.0%; Pred. No. 4e-247;
Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2462 TGCAGAGGTCAGATATCTTCAGGGTGTCTATTGCTGGGGCCCGCCAGTCCTCTGATCT 2521
DB 493 TGCAGAGGTCAGATATCTTCAGGGTGTCTATTGCTGGGGCCCGCCAGTCCTCTGATCT 434
QY 2522 TCTATGATACAGGATACACGGAACTTATATGGTTCACCTCGACCAAGATGAACAGGGCT 2581
DB 433 TCTATGATACAGGATACACGGAACTTATATGGTTCACCTCGACCAAGATGAACAGGGCT 374
QY 2582 ATTACTTACGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAATGCTTTAC 2641
DB 373 ATTACTTACGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAATGCTTTAC 314
QY 2642 TGCTCTTACATGGTTTCTCGGATGAGATGTCCATTTTGGCACATACAGATATATTCTGA 2701
DB 313 TGCTCTTACATGGTTTCTCGGATGAGATGTCCATTTTGGCACATACAGATATATTCTGA 254

```

```

QY 2702 GTTTTGTAGTGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGACACA 2761
Db 253 GTTTTGTAGTGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGACACA 194
QY 2762 GCATAAGAGTTCTCTGAATCGGAGAGAACATATGAGCTGATCTTTGGCATCTACCTTCAAG 2821
Db 193 GCATAAGAGTTCTCTGAATCGGAGAGAACATATGAGCTGATCTTTGGCATCTACCTTCAAG 134
QY 2822 AAAACCTTGGATCAGCTATTGCTCTCTAAAGCTGATATATTTTGGACCTGCTGTAGAAT 2881
Db 133 AAAACCTTGGATCAGCTATTGCTCTCTAAAGCTGATATATTTTGGACCTGCTGTAGAAT 74
QY 2882 CTCTGGTATACACTGGCTATTATTAACCAATGAGGAGTTTATCAACAGAGAAACACAGA 2940
Db 73 CTCTGGTATACACTGGCTATTATTAACCAATGAGGAGTTTATCAACAGAGAAACACAGA 15

RESULT 116
BE799580
LOCUS 601592450F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3946333 5',
DEFINITION mRNA sequence.
ACCESSION BE799580
VERSION BE799580.1 GI:10220778
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 661)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW806 row: d column: 14
High quality sequence start: 58
High quality sequence stop: 659.
Location/Qualifiers
1..661
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3946333"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_7"
/notes="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

```

Query Match 15.3%; Score 478; DB 7; Length 661;
Best Local Similarity 99.8%; Pred. No. 1.4e-246;
Matches 528; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 648 TGGAAACAGTCGGAAATGCTCTACGATATACCAAGGAAGTGGAACTTTCTGTTTCA 707
Db 133 TGGAAACAGTCGGAAATGCTCTACGATATACCAAGGAAGTGGAACTTTCTGTTTCA 192
```

```

QY 708 AGCCGCTAGTGGAAATTTATCAGTAAAAAGATGGAGGGCCACAAGGATTTTACGCAACAAC 767
Db 193 AGCCGCTAGTGGAAATTTATCAGTAAAAAGATGGAGGGCCACAAGGATTTTACGCAACAAC 252
QY 768 TTTAAGGCCCAATCTAGTGGAAACTAGTTGTGCCAATACGGAATGATCCAAAATTTATG 827
Db 253 TTTAAGGCCCAATCTAGTGGAAACTAGTTGTGCCAATACGGAATGATCCAAAATTTATG 312
QY 828 CCCGCTGATCCAGACTCGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACAT 887
Db 313 CCCGCTGATCCAGACTCGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACAT 372
QY 888 GGTACACGAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAGA 947
Db 373 GGTACACGAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAGA 432
QY 948 AGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTC 1007
Db 433 AGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTC 492
QY 1008 TGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCACTGGTGGTAAATTTCTTAGAT 1067
Db 493 TGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCACTGGTGGTAAATTTCTTAGAT 552
QY 1068 TCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTACATCCCTATGTT 1127
Db 553 TCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTACATCCCTATGTT 612
QY 1128 GGAACAAAGAGGGGAGAGATTCATTCGTTATCTTAAACAGGTACAGCA 1176
Db 613 GGAACAAAGAGGGGAGAGATTCATTCGTTATCTTAAACAGGTACAGCA 661

RESULT 117
AI243857
LOCUS 577 bp mRNA linear EST 01-DEC-1998
DEFINITION Q89F09.x1 Soares NFL T GBC SJ Homo sapiens cDNA clone
IMAGE:1854185 3', similar to WP.K02F2.1 CEI7153 PEPTIDASE ;, mRNA
sequence.
ACCESSION AI243857
VERSION AI243857.1 GI:3839254
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 577)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1276 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 435.
Location/Qualifiers
1..577
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1854185"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC SJ"
/notes="Organ: pooled; Vector: pT7T3D-PacI; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (fetal lung NbHL19W, testis NHT, and
B-cell NCI CGAP GCB1) were mixed, and 88 circles were made
in vitro. Following HAP purification, this DNA was used as
```


tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN		Query Match	15.3%; Score 476; DB 1; Length 577;	
		Best Local Similarity	100.0%; Pred. No. 1.6e-245;	
		Matches	476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	2634	TGCTTACTGCTCTTACATGTTTCTCGATGAGATGTCATTTTGGACATACAGTAT	2693	
DB	476	TGCTTACTGCTCTTACATGTTTCTCGATGAGATGTCATTTTGGACATACAGTAT	417	
QY	2694	ATTACTGAGTTTTTTTGTAGGGCTGGAAAGCCATATGATTTACAGATCTATCCTCAGGA	2753	
DB	416	ATTACTGAGTTTTTTTGTAGGGCTGGAAAGCCATATGATTTACAGATCTATCCTCAGGA	357	
QY	2754	GAGACACAGCATAGAGTTCTGATCGGAGACATTTATGAATGCTATCTTTTGCACTA	2813	
DB	356	GAGACACAGCATAGAGTTCTGATCGGAGACATTTATGAATGCTATCTTTTGCACTA	297	
QY	2814	CCTTCAAGAAACCTTGGATCACGTTATGCTCTTAAAGTGTATATAATTTGACCTGT	2873	
DB	296	CCTTCAAGAAACCTTGGATCACGTTATGCTCTTAAAGTGTATATAATTTGACCTGT	237	
QY	2874	GTAGAACTCTCTGGTATACATGCTGCTTAAAGTGTATATAATTTGACCTGT	2933	
DB	236	GTAGAACTCTCTGGTATACATGCTGCTTAAAGTGTATATAATTTGACCTGT	177	
QY	2934	ACAGAAATGATCATCATTTTGCATCTGCTGCAATGTAACATCTACTCTGAAATATAA	2993	
DB	176	ACAGAAATGATCATCATTTTGCATCTGCTGCAATGTAACATCTACTCTGAAATATAA	117	
QY	2994	TGTGGTGCCATGACAGGGGCTACGGTTTGTGGTGTATCTTAATACCTTACCCCATG	3053	
DB	116	TGTGGTGCCATGACAGGGGCTACGGTTTGTGGTGTATCTTAATACCTTACCCCATG	57	
QY	3054	CTCAAAATCAATGATACATATCTCTGAGAGACCCAGCAATACCAATGAATTAAT	3109	
DB	56	CTCAAAATCAATGATACATATCTCTGAGAGACCCAGCAATACCAATGAATTAAT	1	

RESULT 118	DA798581	575 bp	mRNA	linear	EST 09-NOV-2005
LOCUS	DA798581	OCBBF3	Homo sapiens	cDNA clone	OCBBF3002013 5', mRNA
DEFINITION	sequence.				
ACCESSION	DA798581				
VERSION	DA798581.1	GI:81236986			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1. (Bases 1 to 575) Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushiya, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.				
TITLE	Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes				
JOURNAL	Genome Res. 16 (1), 55-65 (2006)				
PUBLISHED	16344560				
COMMENT	Contact: Takao Isogai FLJ Project (HRI Team) Helix Research Institute				

2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.
Location: RAB.
Location/Qualifiers
1. 575
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OCBBF3002013"
/tissue_type="brain"
/dev_stage="fetal"
/clone_lib="OCBBF3"
/note="Vector: pME18SFL3"

ORIGIN

Query Match		15.1%; Score 472; DB 9; Length 575;	
Best Local Similarity		99.7%; Pred. No. 2.4e-243;	
Matches		572; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	8	AAAGCTCCGAGCCCAAGCCGCTCTACTGCGCGCGCTCTTCTTGTAGTCCGCTTCGC	67
DB	2	AAAGCTCCGAGCCCAAGCCGCTCTACTGCGCGCGCTCTTCTTGTAGTCCGCTTCGC	61
QY	68	CGCTCGGTGTTCACCGCGCGCGCGCGAGGAGCCACTGCAACCCAGGACCGAGTGA	127
DB	62	CGCTCGGTGTTCACCGCGCGCGCGCGAGGAGCCACTGCAACCCAGGACCGAGTGA	121
QY	128	GGCGGCGCAGCATGAAGCGCGCGCGCTCTAGCGCAGCTCGCGAGCGTCCGGGC	187
DB	122	GGCGGCGCAGCATGAAGCGCGCGCGCTCTAGCGCAGCTCGCGAGCGTCCGGGC	181
QY	188	GGCGGCGGAGGAGGAAATGCAATGCGCAGCAGCAATGAAACAGACAGCTGGTG	247
DB	182	GGCGGCGGAGGAGGAAATGCAATGCGCAGCAGCAATGAAACAGACAGCTGGTG	241
QY	248	TTGAGATATTGAAACCTGCGGACTGTGAGGAGATATTGAATCAGAGATCGGCTAAAT	307
DB	242	TTGAGATATTGAAACCTGCGGACTGTGAGGAGATATTGAATCAGAGATCGGCTAAAT	301
QY	308	TGGAGCCTTTTATGTTGAGCGGTATTCTCTGGAGTCTAGCTTAAAGAGCTGTTGCCGATA	367
DB	302	TGGAGCCTTTTATGTTGAGCGGTATTCTCTGGAGTCTAGCTTAAAGAGCTGTTGCCGATA	361
QY	368	CCAGAAAATATCATGGCTACATGATGGCTAAGGACCAACATGATTTTCATGTTGTGAAGA	427
DB	362	CCAGAAAATATCATGGCTACATGATGGCTAAGGACCAACATGATTTTCATGTTGTGAAGA	421
QY	428	GGAATGATCCAGATGGACCTCAATTCACAGAGATCTATTACCTTGGCATGCTCTGGTGAGA	487
DB	422	GGAATGATCCAGATGGACCTCAATTCACAGAGATCTATTACCTTGGCATGCTCTGGTGAGA	481
QY	488	ACAGAGAAAATACACTGTTTATTTCTGAAATTTCCAAAACATCAATAGAGAGAGCTCT	547
DB	482	ACAGAGAAAATACACTGTTTATTTCTGAAATTTCCAAAACATCAATAGAGAGAGCTCT	541
QY	548	TAATGCTCTCTTGGAGCCTCTTTTGGATCTTTT	581
DB	542	TAATGCTCTCTTGGAGCCTCTTTTGGATCTTTT	575

RESULT 119	
LOCUS	
DEFINITION	
ACCESSION	

DA189698 521 bp mRNA linear EST 28-OCT-2005
DA189698 BRAMY3 Homo sapiens cDNA clone BRAMY3011692 5', mRNA
sequence.
DA189698

VERSION
 DA189698.1 GI:78246420
 EST.
 Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 521)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Oka, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
 Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H.,
 Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Identification of Transcriptional Modulation: Large-scale
 Diversification and Characterization of Putative Alternative
 Promoters of Human Genes
 Genomics Res. 16 (1), 55-65 (2006)
 16344560
 Contact: Takao Isogai
 HRI Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: filj-cdn@infity.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction:
 Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing: RAB.
 Location/Qualifiers
 1..521
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRAMY3011692"
 /tissue_type="amygdala"
 /clone_lib="BRAMY3"
 /note="Vector: pME18SFL3"
 ORIGIN
 Query Match 15.0%; Score 468; DB 9; Length 521;
 Best Local Similarity 99.8%; Pred No. 3.6e-241;
 Matches 518; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AAGTGTAAAGCTCCGAGGCCAGCGCGCTGCTACTGCGCGCGCTCTTTAGTGGCG 60
 DB 3 AAGTGTAAAGCTTCGAGGCCAGCGCGCTGCTACTGCGCGCGCTCTTTAGTGGCG 62
 QY 61 CGTTCCGCGCTGGTGTGTACCGCGCGCGCGCGCGAGGAGCACTGCAACAGGACCG 120
 DB 63 CGTTCCGCGCTGGTGTGTACCGCGCGCGCGCGAGGAGCACTGCAACAGGACCG 122
 QY 121 GAGTGGAGCGCGGAGCATGAGCGCGCGCGCGCGCTCATAGCGCACGTGCGGACCG 180
 DB 123 GAGTGGAGCGCGGAGCATGAGCGCGCGCGCGCGCTCATAGCGCACGTGCGGACCG 182
 QY 181 TCGGCGCGCGCGCGGAGGAGAAATGCAACATGCGCAGCAATGGAACAGAACAG 240
 DB 183 TCGGCGCGCGCGCGGAGGAGAAATGCAACATGCGCAGCAATGGAACAGAACAG 242
 QY 241 CTGGGTGTGAGATATTTGAACTCGGACTGTGAGGAGAAATTTGAAATCAGAGTCCG 300
 DB 243 CTGGGTGTGAGATATTTGAACTCGGACTGTGAGGAGAAATTTGAAATCAGAGTCCG 302
 QY 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGCTGTT 360
 DB 303 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGCTGTT 362
 QY 361 GCCGATACCAAGAAATATCATGGCTACATGCTAAGGCAACCATGATTTTCATGTT 420

Db 363 GCCGATACCAAGAAATATCATGGCTACATGATGCTAAGGCAACCATGATTTTCATGTT 422
 QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCACAGAGATCTTATACCTTGCATGCT 480
 DB 423 GTGAAGAGGAATGATCCAGATGGACCTCATTCACAGAGATCTTATACCTTGCATGCT 482
 QY 481 GGTGAGAACAGAGAAATACACTGTTTATCTCGAATTT 519
 DB 483 GGTGAGAACAGAGAAATACACTGTTTATCTCGAATTT 521
 RESULT 120
 LOCUS AA306656/c
 DEFINITION EST177562 Jurkat T-cells VI Homo sapiens cDNA 5' end, mRNA
 VERSION AA306656
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 457)
 Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
 Buit, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,
 White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C.,
 Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
 Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S.,
 Glodok, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
 Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
 Moreno-Palanco, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
 Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,
 Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
 Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
 Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A.,
 He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
 Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weisner, P., S. Olsen, H.,
 Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Olsen, H.,
 Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,
 Fraser, C.M. and Venter, J.C.
 Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 7566098
 Other ESTs: THC184894
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018659056
 Fax: 3018659423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M3 Reverse
 Location/Qualifiers
 1..457
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="ARCC (inhost):160827"
 /db_xref="taxon:9606"
 /cell_type="T-lymphocyte"
 /clone_lib="Jurkat T-cells VI"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"
 ORIGIN
 Query Match 14.6%; Score 457; DB 1; Length 457;
 Best Local Similarity 100.0%; Pred. No. 3.3e-235;
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2658	CCTGGATGAGAAATGTCATTTTGACATACCAAGTATATTTAGTGAGGCG	2717
Db	457	CCTGGATGAGAAATGTCATTTTGACATACCAAGTATATTTAGTGAGGCG	398
QY	2718	TGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGATAAGATTCCTGA	2777
Db	397	TGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGATAAGATTCCTGA	338
QY	2778	ATCGGAGAACATTTATGAATGTCATCTTTTGCACTACCTTTCAAGAAAACCTTGGATCAG	2837
Db	337	ATCGGAGAACATTTATGAATGTCATCTTTTGCACTACCTTTCAAGAAAACCTTGGATCAG	278
QY	2838	TATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGG	2897
Db	277	TATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGG	218
QY	2898	CTATTTAAACCAATGAGGAGGTTTAATCAACAGAAAACACAGAAATGATCATCATATTTT	2957
Db	217	CTATTTAAACCAATGAGGAGGTTTAATCAACAGAAAACACAGAAATGATCATCATATTTT	158
QY	2958	GATACCTGCCATGTAACATCTCTCTGAAAATAAATGTGGCCATGACAGGGGCTTACG	3017
Db	157	GATACCTGCCATGTAACATCTCTCTGAAAATAAATGTGGCCATGACAGGGGCTTACG	98
QY	3018	GTTTGTGTAGTAACTTAATACCTTAACCCCATCTCAAAATCAAAATGATACATATTC	3077
Db	97	GTTTGTGTAGTAACTTAATACCTTAACCCCATCTCAAAATCAAAATGATACATATTC	38
QY	3078	CTGAGAGACCCAGCAATACCAATGAAGAAATTAATAAAA 3114	
Db	37	CTGAGAGACCCAGCAATACCAATGAAGAAATTAATAAAA 1	
RESULT 121			
DB156905			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
COMMENT			
FEATURES			

source			
1..576			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="THYMU3035791"			
/tissue_type="thymus"			
/clone_lib="THYMU3"			
/note="Vector: pME18SFL3"			
ORIGIN			
Query Match			
Best Local Similarity 99.8%; Score 455; DB 9; Length 576;			
Matches 575; Conservative 0; Mismatches 0; Indels 1; Gaps 1;			
QY	1	AAGTGCTAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCCGCTGCTTCTTAGTGCGG	60
Db	1	AAGTGCTAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCCGCTGCTTCTTAGTGCGG	60
QY	61	CGTTGCGCGCCCTGGGTTGTCAACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGGACG	120
Db	61	CGTTGCGCGCCCTGGGTTGTCAACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGGACG	120
QY	121	GAGTGAGGCGCGCGCAGCATGAAGC-GGCGCAGCGCGCGCTCCATAGCGCAGCTGCGGACG	179
Db	121	GAGTGAGGCGCGCGCAGCATGAAGC-GGCGCAGCGCGCGCTCCATAGCGCAGCTGCGGACG	180
QY	180	GTCCGCGCGCGCGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAGAACA	239
Db	181	GTCCGCGCGCGCGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAGAACA	240
QY	240	GCTGGGCTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAAGATCG	299
Db	241	GCTGGGCTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAAGATCG	300
QY	300	GCCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCT	359
Db	301	GCCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCT	360
QY	360	TGCCGATACACAGAAATATCATGCTACATGATGGCTAAGGCCACCATGATTTTCATGTT	419
Db	361	TGCCGATACACAGAAATATCATGCTACATGATGGCTAAGGCCACCATGATTTTCATGTT	420
QY	420	TGTGAAGAGGAATGATCCAGATGACCTCATTCAGACAGATCTTATACCTTGGCATGTC	479
Db	421	TGTGAAGAGGAATGATCCAGATGACCTCATTCAGACAGATCTTATACCTTGGCATGTC	480
QY	480	TGCTGAGAGACAGAGAAAATACACTGTTTTTATCTGAAATTCCTCAAAATTCATCATAGAGC	539
Db	481	TGCTGAGAGACAGAGAAAATACACTGTTTTTATCTGAAATTCCTCAAAATTCATCATAGAGC	540
QY	540	AGCAGCTTTTAATGCTCTCTTTGGAAGCCCTCTTTTGA 575	
Db	541	AGCAGCTTTTAATGCTCTCTTTGGAAGCCCTCTTTTGA 576	
RESULT 122			
DA607747			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			

14.6%; Score 455; DB 9; Length 576;
Best Local Similarity 99.8%; Pred. No. 4e-234;
Matches 575; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 AAGTGCTAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCCGCTGCTTCTTAGTGCGG 60
1 AAGTGCTAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCCGCTGCTTCTTAGTGCGG 60
61 CGTTGCGCGCCCTGGGTTGTCAACCGCGCGCGCGAGGAGCCACTGCAACACGAGGACG 120
61 CGTTGCGCGCCCTGGGTTGTCAACCGCGCGCGCGAGGAGCCACTGCAACACGAGGACG 120
121 GAGTGAGGCGCGCGCAGCATGAAGC-GGCGCAGCGCGCGCTCCATAGCGCAGCTGCGGACG 179
121 GAGTGAGGCGCGCGCAGCATGAAGC-GGCGCAGCGCGCGCTCCATAGCGCAGCTGCGGACG 180
180 GTCCGCGCGCGCGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAGAACA 239
181 GTCCGCGCGCGCGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAGAACA 240
240 GCTGGGCTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAAGATCG 299
241 GCTGGGCTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAAGATCG 300
300 GCCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCT 359
301 GCCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCT 360
360 TGCCGATACACAGAAATATCATGCTACATGATGGCTAAGGCCACCATGATTTTCATGTT 419
361 TGCCGATACACAGAAATATCATGCTACATGATGGCTAAGGCCACCATGATTTTCATGTT 420
420 TGTGAAGAGGAATGATCCAGATGACCTCATTCAGACAGATCTTATACCTTGGCATGTC 479
421 TGTGAAGAGGAATGATCCAGATGACCTCATTCAGACAGATCTTATACCTTGGCATGTC 480
480 TGCTGAGAGACAGAGAAAATACACTGTTTTTATCTGAAATTCCTCAAAATTCATCATAGAGC 539
481 TGCTGAGAGACAGAGAAAATACACTGTTTTTATCTGAAATTCCTCAAAATTCATCATAGAGC 540
540 AGCAGCTTTTAATGCTCTCTTTGGAAGCCCTCTTTTGA 575
541 AGCAGCTTTTAATGCTCTCTTTGGAAGCCCTCTTTTGA 576

DA607747 IMR322 Homo sapiens cDNA clone IMR322005026 5', mRNA
sequence.
DA607747
EST. DA607747.1 GI:80534653
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 587)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers

Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T., and Sugano, S. Identification of Transcriptional Modulation: Large-scale Promoters of Human Genes

JOURNAL PUBMED COMMENT

Genome Res. 16 (1), 55-65 (2006)

16344560

Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com

NEO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES source

Location/Qualifiers

1. 587

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMR322005026"

/cell_type="neuroblastoma"

/cell_line="IMR322"

/clone_lib="IMR322"

/notes="Vector: pME18SFL3"

ORIGIN

Query Match 14.5%; Score 452; DB 9; Length 587;
Best Local Similarity 99.8%; Pred. No. 1.7e-232;
Matches 572; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AAGTGTAAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTCTTCTTAGTGGCG 60

Db 2 AAGTGTAAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTCTTCTTAGTGGCG 61

QY 61 CGTTCCCGCCCTGGTGTGTCACCGCGCGCGCGGAGGAGCCACTGCACACGAGCG 120

Db 62 CGTTCCCGCCCTGGTGTGTCACCGCGCGCGCGGAGGAGCCACTGCACACGAGCG 121

QY 121 GAGTGGAGCGCGCAGCATGAAGCGCGGAGCGCGCGCTCATAGCGCAC -GTCGGGAGC 179

Db 122 GAGTGGAGCGCGCAGCATGAAGCGCGGAGCGCGCGCTCATAGCGCACGTCGGGAGC 181

QY 180 GTCCGGCGCGCGCGCGGAGGAGGAAATGCACATGCGACAGCAATGGAACAGAAC 239

Db 182 GTCCGGCGCGCGCGGAGGAGGAAATGCACATGCGACAGCAATGGAACAGAAC 241

QY 240 GCTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGGATCG 299

Db 242 GCTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGGATCG 301

QY 300 GCTTAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGATGCT 359

Db 302 GCTTAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGATGCT 361

QY 360 TCCCATACCAAGAAATATCATGCTCATGATGGCTAAGGACCAACATGATTTTCATGTT 419

Db 362 TCCCATACCAAGAAATATCATGCTCATGATGGCTAAGGACCAACATGATTTTCATGTT 421

QY 420 TGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTCTTGGCATGTC 479

Db 422 TGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTCTTGGCATGTC 481

QY 480 TGGTGAGAACAGAGAAATACACTGTTTTTATTTCTGAAATTTCCCAAAATATCAATAGAGC 539

Db 482 TGGTGAGAACAGAGAAATACACTGTTTTTATTTCTGAAATTTCCCAAAATATCAATAGAGC 541

QY 540 AGCAGTCTTAATGCTCTCTTGGAGGCTCTTTT 572

Db 542 AGCAGTCTTAATGCTCTTGGAGGCTCTTTT 574

RESULT 123

DA145947

LOCUS

DA145947 BRAMY2 Homo sapiens cDNA clone BAMY2002583 5', mRNA

DEFINITION

sequence.

DA145947

ACCESSION

DA145947.1 GI:79192295

VERSION

EST.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1. (bases 1 to 570)

AUTHORS

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaquri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Sugano, S., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T., and Sugano, S. Identification of Transcriptional Modulation: Large-scale Promoters of Human Genes

TITLE

Genome Res. 16 (1), 55-65 (2006)

JOURNAL PUBMED COMMENT

16344560

Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com

NEO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES source

Location/Qualifiers

1. 570

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRAMY2002583"

/tissue_type="amygdala"

/clone_lib="BRAMY2"

/note="Vector: pME18SFL3"

ORIGIN

Query Match 14.4%; Score 449; DB 9; Length 570;
Best Local Similarity 99.8%; Pred. No. 7.2e-231;
Matches 569; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2 AGTGTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTCTTCTTAGTGGCG 61

Db 1 AGTGTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTCTTCTTAGTGGCG 60

QY 62 GTTCCGCGCTGGGTGTGTCACCGCGCGCGCGGAGGAGCCACTGCAACAGGACCG 121

Db 61 GTTCCGCGCTGGGTGTGTCACCGCGCGCGCGGAGGAGCCACTGCAACAGGACCG 120

QY 122 AGTGGAGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAC -GTCGGGACCG 180

Db 121 AGTGGAGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCACGTCGGGACCG 180

QY 181 TCCGGCGCGCGCGCGGAGGAGAAATGCAACATGGCAGCAGCAATGGAACAGAACAG 240

Db 181 TCCGGCGCGCGCGCGGAGGAGAAATGCAACATGGCAGCAGCAATGGAACAGAACAG 240

[illegible]

RESULT	124
DB235712	
LOCUS	DB235712 TRACH3 Homo sapiens cDNA clone TRACH3030760 5', mRNA
DEFINITION	DB235712 linear EST 04-DEC-2005
	570 bp mRNA
ACCESSION	DB235712
VERSION	DB235712.1 GI:83202774
KEYWORDS	sequence.
SOURCE	DB235712 Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
	Hominidae; Homo.

1 (bases 1 to 570)

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Kimura, K., R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Itie, R., Kushida, N., Yoneyama, T., Otsuka, K., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, K., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T., and Sugano, S. Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes

Promoter Res. 16 (1), 55-65 (2006)

JOURNAL
PUBLISHED
COMMENT

Genome Res. 16 (1), 55-65 (2006)

Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.

```

FEATURES
source
Location/Qualifiers
1. 570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TRACH3030760"
/tissue_type="trachea"
/clone_lib="TRACH3"
/note="Vector: pME18SFL3"

```

ORIGIN

Query Match	14.4%;	Score 449;	DB 9;	Length 570;
Best Local Similarity	99.8%;	Pred. Nism.7.2e-231;		
Matches 569;	Conservative	0;	Indels 1;	Gaps 1;
8	AAAGCTCGAGGCCAAGCCGCTGCTACTACTCGCCGCTGCTCTTAGTGC CGCGTTCCG 67			
1	AAAGCTCGAGGCCAAGCCGCTGCTACTCGCCGCTGCTCTTAGTGC CGCGTTCCG 60			
68	CGCTCGGTTGTTCACCGGCGCCGCGCGAGGAAGCCATCTCAACACGAGACCGAGTGG 127			
61	CGCTCGGTTGTTCACCGGCGCCGCGCGAGGAAGCCATCTCAACACGAGACCGAGTGG 120			
128	GGCGGCGCAGCATGAAGC--GGCGCAGGCCGCTCCATAGCCGACGCTCGGACGCGTCCGGG 186			
121	GGCGGCGCAGCATGAAGCAGCGGCGCAGGCCGCTCCATAGCCGACGCTCGGACGCGTCCGGG 180			
187	CGGGCGCGGGGGAAGGAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTGGGT 246			
181	CGGGCGCGGGGGAAGGAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTGGGT 240			
247	GTTGAGATATTGGAACCTGGGACTGTGAGGAGATATTGAATCACAGATCGGCGCTAAA 306			
241	GTTGAGATATTGGAACCTGGGACTGTGAGGAGATATTGAATCACAGATCGGCGCTAAA 300			
307	TTGAGCGCTTTTATGTTTGAGCGGTATTCCTGGAGTCAGCTTTAAAAGTGCTTGC CGAT 366			
301	TTGAGCGCTTTTATGTTTGAGCGGTATTCCTGGAGTCAGCTTTAAAAGTGCTTGC CGAT 360			
367	ACCGAAATAATCATGCTCATGATGCTTAAGCCACCATGATTTCATGTTTGTGAAG 426			
361	ACCGAAATAATCATGCTCATGATGCTTAAGCCACCATGATTTCATGTTTGTGAAG 420			
427	AGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTCCCATGTCTGTGAG 486			
421	AGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTCCCATGTCTGTGAG 480			
487	AAACAGAGAAAATACACTGTGTTTATCTTGAAATTCGAAACTTCAATCAGACAGCAGTC 546			
481	AAACAGAGAAAATACACTGTGTTTATCTTGAAATTCGAAACTTCAATCAGACAGCAGTC 540			
547	TTAATGCTCTCTTGGAGCGCTCTTTTGGAT 576			
541	TTAATGCTCTCTTGGAGCGCTCTTTTGGAT 570			

[illegible]

REFERENCE AUTHORS	JOURNAL PUBMED	TITLE	COMMENT
1. (bases 1 to 568)			
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Teurittan, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushiida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibaishi, T., Takahashi-Fujii, A., Sugano, S., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isoigai, F. and Sugano, S.		Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes	
	Genome Res. 16 (1), 55-65 (2006)		
	16344560		
	Contact: Takao Isoigai		
	FLJ Project (HRI Team)		

Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdnaenifty.com

NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.

FEATURES
source
Location/Qualifiers
1..568
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRACE2036520"
/tissue_type="cerebellum"
/clone_lib="BRACE2"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 14.3%; Score 447; DB 9; Length 568;
Best Local Similarity 99.8%; Pred. No. 8.7e-230;
Matches 567; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 AAGTGCTAAAGCCTCCGAGCCCAAGCCGCTGCTACTGCGCCGCTGCTTCTTAGTGCCG 60
Db 1 AAGTGCTAAAGCCTCCGAGCCCAAGCCGCTGCTACTGCGCCGCTGCTTCTTAGTGCCG 60
QY 61 CGTTCGCGCCTGGTGTTCACCGCGCCGCGCGAGGAAAGCCACTGCAACCAAGGACCG 120
Db 61 CGTTCGCGCCTGGTGTTCACCGCGCCGCGCGAGGAAAGCCACTGCAACCAAGGACCG 120
QY 121 GAGTGGAGCGCGCAGCATGAGCGCGCAGCGCCGCTCCATAGC-GCACGTGCGGACG 179
Db 121 GAGTGGAGCGCGCAGCATGAGCGCGCAGCGCCGCTCCATAGC-GCACGTGCGGACG 180
QY 180 GTCCGCGCGCGCGCGGAGAAAGGAAATGCAACATGCGCAGCAGCAATGGAACAGACA 239
Db 181 GTCCGCGCGCGCGCGGAGAAAGGAAATGCAACATGCGCAGCAGCAATGGAACAGACA 240
QY 240 GCTGGGTGTGAGATATTGAACTGCGGACTGTGAGAGAGAAATGGAATCAGAGATCG 299
Db 241 GCTGGGTGTGAGATATTGAACTGCGGACTGTGAGAGAGAAATGGAATCAGAGATCG 300
QY 300 GCTTAAATTGGAGCCTTTTATGTGTAGCGGTATTCCTGGAGTCAGCTTAAAGAGTGT 359
Db 301 GCTTAAATTGGAGCCTTTTATGTGTAGCGGTATTCCTGGAGTCAGCTTAAAGAGTGT 360
QY 360 TGCCGATACCAAGAAATATCATGCTACATGATGGCTTAAGCAGCACCATGTTTATGTT 419
Db 361 TGCCGATACCAAGAAATATCATGCTACATGATGGCTTAAGCAGCACCATGTTTATGTT 420
QY 420 TGTGAAGAGGAATGATCCAGATGACCTCATTGAGCAGATCTATTACCTTGCATGTC 479
Db 421 TGTGAAGAGGAATGATCCAGATGACCTCATTGAGCAGATCTATTACCTTGCATGTC 480
QY 480 TGGTGAGAAACAGAGAAATACACTGTTTATCTGAAATCCCAAACTATCAATAGAGC 539
Db 481 TGGTGAGAAACAGAGAAATACACTGTTTATCTGAAATCCCAAACTATCAATAGAGC 540
QY 540 AGCAGTCTTAATGCTCTCTGGAGCCT 567
Db 541 AGCAGTCTTAATGCTCTCTGGAGCCT 568

RESULT 126
DB161177
LOCUS
DEFINITION
DB161177 THYMUS3 Homo sapiens cDNA clone THYMUS041608 5', mRNA
sequence.
ACCESSION
DB161177

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT

DB161177.1 GI:83155018
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 570)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
Yoneyama, T., Otsuka, K., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishidashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdnaenifty.com

NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.

FEATURES
source
Location/Qualifiers
1..570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="THYMUS041608"
/tissue_type="thymus"
/clone_lib="THYMUS3"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 14.3%; Score 447; DB 9; Length 570;
Best Local Similarity 99.8%; Pred. No. 8.7e-230;
Matches 567; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 AAGTGCTAAAGCCTCCGAGCCCAAGCCGCTGCTACTGCGCCGCTGCTTCTTAGTGCCG 60
Db 3 AAGTGCTAAAGCCTCCGAGCCCAAGCCGCTGCTACTGCGCCGCTGCTTCTTAGTGCCG 62
QY 61 CGTTCGCGCCTGGTGTTCACCGCGCGCGCGAGGAAAGCCACTGCAACCAAGGACCG 120
Db 63 CGTTCGCGCCTGGTGTTCACCGCGCGCGCGAGGAAAGCCACTGCAACCAAGGACCG 122
QY 121 GAGTGGAGCGCGCAGCATGAGCGCGCAGCGCCGCTCCATAGCAGCAGTGGAGCAGG 180
Db 123 GAGTGGAGCGCGCAGCATGAGCGCGCAGCGCCGCTCCATAGCAGCAGTGGAGCAGG 182
QY 181 TCCTGGGCGCGCGCGGAGGAAATGCAATGCGCAGCAGCAATGGAACAGAACAG 240
Db 183 TCCTGGGCGCGCGCGGAGGAAATGCAATGCGCAGCAGCAATGGAACAGAACAG 242
QY 241 CTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCG 300
Db 243 CTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCG 302
QY 301 CCTAAATTGGAGCC-TTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGTGTCT 359
Db 303 CCTAAATTGGAGCC-TTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGTGTCT 362
QY 360 TGCCGATACCAAGAAATATCATGCTACATGATGGCTAAGGCGCACCATGTTTATGTT 419

```

Db      363  TGCGATACAGAAATATCATGGCTACATGATGGCTAAGGCACCATGATTTTCATGTT 422
Qy      420  TGTGAAGAGAAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTGGCATGTC 479
Db      423  TGTGAAGAGAAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTGGCATGTC 482
Qy      480  TGGTGAGAACAGAGAAATACACTGTTTTATTCTGAAATTCCTCAAAACTATCAATAGAGC 539
Db      483  TGGTGAGAACAGAGAAATACACTGTTTTATTCTGAAATTCCTCAAAACTATCAATAGAGC 542
Qy      540  AGCAGTCTTAATGCTCTCTTTGGAGGCT 567
Db      543  AGCAGTCTTAATGCTCTCTTTGGAGGCT 570

RESULT 127
AA465309
LOCUS   aa24b08.t1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:814167 5',
DEFINITION
mRNA sequence.
ACCESSION
VERSION AA465309.1 GI:2191476
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 469)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham.

FEATURES
         source
         1. 469
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="GDB:6032116"
            /db_xref="taxon:9606"
            /clone="IMAGE:814167"
            /tissue_type="germinal center 'B cell'"
            /lab_host="DH10B"
            /clone_lib="NCI_CGAP_GCB1"
            /note="Vector: p7T3D-Paci; Site_1: Not I; Site_2: Eco RI;
            1st strand cDNA was prepared from human tonsillar cells
            enriched for germinal center B cells by flow sorting
            (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
            David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA
            synthesis was primed with a Not I - oligo(dT) primer
            [5'-TGTTACCAATCTGAATGCGAGCGGCTCATTTTCTTTTCTTTT-3',
            ]. Double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified p7T3 vector. Library
            went through one round of normalization, and was
            constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 14.3%; Score 445; DB 1; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e-228;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1560  GTTTATTTTGGCTCTGAATGCAGAAACAGGTTTCGGTCATTTATACAAAATTACATCTAT 1619
Db      1  GTTTATTTTGGCTCTGAATGCAGAAACAGGTTTCGGTCATTTATACAAAATTACATCTAT 60
Qy      1620  TTTAAAGGAAACAAATATAACAGTCCAGTGGTGGGCTGCTGCTCCAAAGTGAATTCAA 1679
Db      61  TTTAAAGGAAACAAATATAACAGTCCAGTGGTGGGCTGCTGCTCCAAAGTGAATTCAA 120

```

```

Qy      1680  GTGTCTCTCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTTGGCCGGCA 1739
Db      121  GTGTCTCTCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTTGGCCGGCA 180
Qy      1740  TGGATCTTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1799
Db      181  TGGATCTTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 240
Qy      1800  CTCCTCTTTAGAGCATCACCTGTAGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAG 1859
Db      241  CTCCTCTTTAGAGCATCACCTGTAGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAG 300
Qy      1860  GCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGTGTGACTCTTTAT 1919
Db      301  GCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGTGTGACTCTTTAT 360
Qy      1920  AAGTAAGTATAGTACCAAGAGAGATCCACACTGTGTCCTTTTACAAGCTATCAAGTCC 1979
Db      361  AAGTAAGTATAGTACCAAGAGAGATCCACACTGTGTCCTTTTACAAGCTATCAAGTCC 420
Qy      1980  TGAAGATGACCCAACTTGCAGAAAACA 2004
Db      421  TGAAGATGACCCAACTTGCAGAAAACA 445

AA831859 437 bp mRNA linear EST 18-MAR-1998
OC93hl12.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1357319 3',
mRNA sequence.
AA831859
VERSION AA831859.1 GI:2904958
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 437)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1067 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 399.
         Location/Qualifiers
         1. 437
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:1357319"
            /tissue_type="germinal center B cell"
            /lab_host="DH10B"
            /clone_lib="NCI_CGAP_GCB1"
            /note="Vector: p7T3D-Paci; Site_1: Not I; Site_2: Eco RI;
            1st strand cDNA was prepared from human tonsillar cells
            enriched for germinal center B cells by flow sorting
            (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
            David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA
            synthesis was primed with a Not I - oligo(dT) primer
            [5'-TGTTACCAATCTGAATGCGAGCGGCTCATTTTCTTTTCTTTT-3',
            ]. Double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified p7T3 vector. Library
            went through one round of normalization, and was
            constructed by Bento Soares and M. Fatima Bonaldo."

```

[5'-TGTACCAATCTGAAGTGGGAGCGCGCTCATTTTTTTTTTTTTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTV73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 14.0%; Score 437; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.3e-224;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2683 CATACAGTATATTACTAGTCTTTTGTAGTGGGCTGGAAGCCATATGTTTACAGATC 2742
DB CATACAGTATATTACTAGTCTTTTGTAGTGGGCTGGAAGCCATATGTTTACAGATC 378
QY 2743 TATCTCTCAGGAGACACAGATAGAGTTCCTGAATCGGAGAACTATATGACATGAT 2802
DB TATCTCTCAGGAGACACAGATAGAGTTCCTGAATCGGAGAACTATATGACATGAT 318
QY 2803 CTTTTTGCACTACCTTCAGAAACCTTGGATCAGATATGCTGCTCTAAAGTGATATA 2862
DB CTTTTTGCACTACCTTCAGAAACCTTGGATCAGATATGCTGCTCTAAAGTGATATA 258
QY 2863 TTTTGACCTGTGTAGACTCTCTGTATACACTGGCTATTTAACCAATGAGGAGTTTA 2922
DB TTTTGACCTGTGTAGACTCTCTGTATACACTGGCTATTTAACCAATGAGGAGTTTA 198
QY 2923 ATCACAGAAACACAGATGATCATATTTTGATACCTGCCATGTAACTACTCTC 2982
DB ATCACAGAAACACAGATGATCATATTTTGATACCTGCCATGTAACTACTCTC 138
QY 2983 CTGAATTAATGTGGTGCATGACGGGTCTACGGTTTGTGGTAGTAACTTAATACCTT 3042
DB CTGAATTAATGTGGTGCATGACGGGTCTACGGTTTGTGGTAGTAACTTAATACCTT 78
QY 3043 AACCCCATGCTCAAAATCAATGATACATATTCCTGAGACCCAGCAATACCATAG 3102
DB AACCCCATGCTCAAAATCAATGATACATATTCCTGAGACCCAGCAATACCATAG 18
QY 3103 AATTACTAAAAA 3119
DB AATTACTAAAAA 1

RESULT 129

DA455896 558 bp mRNA linear EST 07-NOV-2005
LOCUS DA455896 CTONG3 Homo sapiens cDNA clone CTONG3000794 5', mRNA
DEFINITION

ACCESSION

DA455896
DA455896.1 GI:81156902

VERSION

EST.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo

1 (bases 1 to 558)

Kimura K., Wakamatsu A., Suzuki Y., Ota T., Nishikawa T.,

Yamashita R., Yamamoto J., Sekine M., Tsuritani K., Wakaguri H.,

Ishii S., Sugiyama T., Saito K., Isono Y., Irie R., Kishida N.,

Yoneyama T., Otsuka R., Kanda K., Yokoi T., Kondo H., Wagatsuma M.,

Murakawa K., Ishida S., Ishibashi T., Takahashi Fujii A.,

Tanabe T., Nagai K., Kikuchi H., Nakai K., Isogai T. and Sugano S.

Diversification of Transcriptional Modulation: Large-scale

Identification and Characterization of Putative Alternative

Promoters of Human Genes

Genome Res. 16 (1), 55-65 (2006)

16344560

Contact: Takao Isogai

FLJ Project (HRI Team)

Helix Research Institute

TITLE

JOURNAL

PUBLISHED

COMMENT

2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction;
Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.

FEATURES

source

Location/Qualifiers
1..558
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG3000794"
/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG3"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 14.0%; Score 437; DB 9; Length 558;
Best Local Similarity 99.8%; Pred. No. 2.3e-224;
Matches 557; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 8 AAAGCTCGAGGCGCAAGCGCGCTGCTACTGCGCGCGCTGCTTTAGTCCCGTTCGC 67
DB 1 AAAGCTCGAGGCGCAAGCGCGCTGCTACTGCGCGCGCTGCTTTAGTCCCGTTCGC 60
QY 68 CGCTCGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCGGAGTGA 127
DB 61 CGCTCGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCGGAGTGA 120
QY 128 GCGCGCGCAGCATGAAGCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCGGAGTGA 186
DB 121 GCGCGCGCAGCATGAAGCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCGGAGTGA 180
QY 187 CGGCGCGCGGAGGAGAAATGCAACATGCGCAGCAGCAATGGAACAGAGCTGGCT 246
DB 181 CGGCGCGCGGAGGAGAAATGCAACATGCGCAGCAGCAATGGAACAGAGCTGGCT 240
QY 247 GTTGAGATATTTGAAACTGCGGACTGTGAGGAGAGATTTGAAATCAGCAGATCGGCT 306
DB 241 GTTGAGATATTTGAAACTGCGGACTGTGAGGAGAGATTTGAAATCAGCAGATCGGCT 300
QY 307 TTGAGAGCTTTTATGTGTAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTCCGAT 366
DB 301 TTGAGAGCTTTTATGTGTAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTCCGAT 360
QY 367 ACCAGAAAATATCATGGCTACATGATGGCTTAAGGACCAACATGATTCATGTTGTGAAG 426
DB 361 ACCAGAAAATATCATGGCTACATGATGGCTTAAGGACCAACATGATTCATGTTGTGAAG 420
QY 427 AGGAATGATCCAGATGGACTCATTTCAGACAGAACTATTACCTTGCATGCTGGTAG 486
DB 421 AGGAATGATCCAGATGGACTCATTTCAGACAGAACTATTACCTTGCATGCTGGTAG 480
QY 487 AACAGAAAATACACTGTTTATTTCTGAAATTTCCAAAATCATCAATAGAGCAGCTC 546
DB 481 AACAGAAAATACACTGTTTATTTCTGAAATTTCCAAAATCATCAATAGAGCAGCTC 540
QY 547 TTAATGCTCTCTTGGAG 564
DB 541 TTAATGCTCTCTTGGAG 558

RESULT 130

CM427231

LOCUS

DEFINITION

CM427231

ACCESSION

CM427231.1

VERSION

KEYWORDS

CM427231 726 bp mRNA linear EST 16-MAY-2004
17000455488465 GRN_EB Homo sapiens cDNA 5', mRNA sequence.

CM427231.1 GI:47414825
EST.

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 726)
AUTHORS	Brandenberger R., Wei H., Zhang S., Lei S., Murga J., Fisk G.J., Li Y., Xu C., Fang R., Guegler K., Rao M.S., Mandalam R., Lebkowski J. and Stanton L.W.
TITLE	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL	Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED	15146197
COMMENT	Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@geron.com Insert Length: 726 Std Error: 0.00.
FEATURES	Location/Qualifiers
source	1..726
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/tissue_type="embryonic stem cells, embryoid bodies derived from H1, H7 and H9 cells"
	/clone_lib="GRN EB"
	/notes="oligo dt primed, full-length enriched cDNA library from embryoid body outgrowths derived from hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."
ORIGIN	
Query Match	13.9%; Score 434; DB 8; Length 726;
Best Local Similarity	99.7%; Pred. No. 1e-222; 1; Indels 1; Gaps 1;
Matches 604; Conservative	0; Mismatches 0;
QY	1859 GCCTGACTGACCGTGGCTACTCATTCTTGTGTCATAGTCAGCACTGTGACTCTTTA 1918
Db	2 GCTGACTGACCGTGGCTACTCATTCTTGTGTCATAGTCAGCACTGTGACTCTTTA 61
QY	1919 TAACTAGTATAGTAAACAGAGATCCACACTGTGTCCCTTACAGCTATCAAGTC 1978
Db	62 TAACTAGTATAGTAAACAGAGATCCACACTGTGTGTCCCTTACAGCTATCAAGTC 121
QY	1979 CTGAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGCAG 2038
Db	122 CTGAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGCAG 181
QY	2039 GTCCCTTTCCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTA 2098
Db	182 GTCCCTTTCCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTA 241
QY	2099 CATTTGATGGGATGCTTACAGCTCATGATCTACAGCTGGAAGAAATATCTTACTG 2158
Db	242 CATTTGATGGGATGCTTACAGCTCATGATCTACAGCTGGAAGAAATATCTTACTG 301
QY	2159 TGTCTGTTTATATAT-GGTGGTCTCAGGTGAGTGGTGGTGAATATATCGTTTAAAGAGTC 2217
Db	302 TGTCTGTTTATATATGGTGGTCTCAGGTGAGTGGTGGTGAATATATCGTTTAAAGAGTC 361
QY	2218 AAGTATTCGGCTTGAATACCTAGCTCTCTAGTTATGTTGGTTAGTGTAGTAGACAAC 2277
Db	362 AAGTATTCGGCTTGAATACCTAGCTCTCTAGTTATGTTGGTTAGTGTAGTAGACAAC 421
QY	2278 AGGGATCTCTGACCCAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAA 2337
Db	422 AGGGATCTCTGACCCAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAA 481
QY	2338 ATAGAAATTCAGCATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTCATT 2397
Db	482 ATAGAAATTCAGCATCAGGTGGAGGATTCCAATATCTAGCTTCTCGATATGATTCATT 541
QY	2398 GACTTAGATCGTGGGATCCACGGCTGGTCTTATGGAGATACCTCTCCCTGATGGCA 2457
Db	542 GACTTAGATCGTGGGATCCACGGCTGGTCTTATGGAGATACCTCTCCCTGATGGCA 601
QY	2458 TTAATG 2463
Db	602 TTAATG 607
RESULT 131	
AA278626/c	
LOCUS	AA278626 494 bp mRNA linear EST 15-AUG-1997
DEFINITION	z578g11.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:703652 3', mRNA sequence.
ACCESSION	AA278626
VERSION	AA278626.1 GI:1919946
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 494)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 817 Std Error: 0.00 Seq primer: -41m13 fwd. ET from Amersham High quality sequence stop: 471.
FEATURES	Location/Qualifiers
source	1..494
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:703652"
	/tissue_type="germinal center B cell"
	/lab_host="DH10B"
	/clone_lib="NCI_CGAP GCB1"
	/note="Vector: p773D-PacI; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dt) primer [5'-TGTACCAATCTGAGTGGGAGCGGCTCATTTTTTTTTTTTTTT-3', 15'-TGTACCAATCTGAGTGGGAGCGGCTCATTTTTTTTTTTTTTT-3', 1. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN	
Query Match	13.9%; Score 433; DB 1; Length 494;
Best Local Similarity	100.0%; Pred. No. 3.4e-222; 0; Indels 0; Gaps 0;
Matches 433; Conservative	0; Mismatches 0;
QY	2677 TTTCACATACCATATATCTAGTTTATGAGGGCTGGAAAGCCATATGATTTA 2736
Db	433 TTTCACATACCATATATCTAGTTTATGAGGGCTGGAAAGCCATATGATTTA 374
QY	2737 CAGATCTATCTCAGGAGACACACATAAGAGTTCTCTGAATCGGGAGACATTATGAA 2796
Db	373 CAGATCTATCTCAGGAGACACACATAAGAGTTCTCTGAATCGGGAGACATTATGAA 314
QY	2797 CTGATCTTTTGGCTACCTTCAAGAAACCTTGGATCAGTATGCTCTCTAAAGTG 2856
Db	314 CTGATCTTTTGGCTACCTTCAAGAAACCTTGGATCAGTATGCTCTCTAAAGTG 2856

Db 313 CTGCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTATTGCTGCTAAAGTG 254
 QY 2857 ATATATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGA 2916
 Db 253 ATATATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGA 194
 QY 2917 GGTATATCAACAGAAAACACAGAAATTCATCATCATTTTGTATACCTGCGCATGAACAT 2976
 Db 193 GGTATATCAACAGAAAACACAGAAATTCATCATCATTTTGTATACCTGCGCATGAACAT 134
 QY 2977 CTACTCTCTGAAAATAAATGTGTGCGCATGCGAGGGGTCTACGGTTTGTGGTAGTAACTAA 3036
 Db 133 CTACTCTCTGAAAATAAATGTGTGCGCATGCGAGGGGTCTACGGTTTGTGGTAGTAACTAA 74
 QY 3037 TACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATAC 3096
 Db 73 TACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATAC 14
 QY 3097 CATAAGAATTACT 3109
 Db 13 CATAAGAATTACT 1

RESULT 132

DB358826/c

LOCUS

DEFINITION DB358826 HEMBB1 Homo sapiens cDNA clone HEMBB1000150 3', mRNA

ACCESSION DB358826

VERSION DB358826.1 GI:83486188

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

1 (bases 1 to 514)

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,

Yamashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaguri, H.,

Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,

Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuna, M.,

Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,

Tanabe, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.

Diversification and Characterization of Putative Alternative

Promoters of Human Genes

Genome Res. 16 (1), 55-65 (2006)

16344560

Contact: Takao Isogai

FLJ Project (HRI Team)

Helix Research Institute

2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: flj-cdna@nifty.com

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix

Research Institute (HRI); cDNA library construction: Department of

Virology, Institute of Medical Science, University of Tokyo, and

HRI.

JOURNAL PUBMED

COMMENT

FEATURES

source

ORIGIN

Query Match 13.8%; Score 430; DB 9; Length 514;

Best Local Similarity 100.0%; Pred. No. 1.4e-220;

Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2680 GCATACACAGTATATTAAGTCTTTAGTGGAGGCTGGAGCCATATGATTACAG 2739
 Db 430 GCATACACAGTATATTAAGTCTTTAGTGGAGGCTGGAGCCATATGATTACAG 371
 QY 2740 ATCTATCCTCAGGAGAGACACAGCATTAAGAGTTCCTGAATCGGAGAACATTATGAAC 2799
 Db 370 ATCTATCCTCAGGAGAGACACAGCATTAAGAGTTCCTGAATCGGAGAACATTATGAAC 311
 QY 2800 CATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACCTATTGCTCTTAAAGTGATA 2859
 Db 310 CATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACCTATTGCTCTTAAAGTGATA 251
 QY 2860 TAATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGCTATTAAACCAATGAGGAGGT 2919
 Db 250 TAATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGCTATTAAACCAATGAGGAGGT 191
 QY 2920 TTAATCAACAGAAAACACAGAAATTCATCATCATTTTGTGATPACCTGCATGTAAACATCTA 2979
 Db 190 TTAATCAACAGAAAACACAGAAATTCATCATCATTTTGTGATPACCTGCATGTAAACATCTA 131
 QY 2980 CTCCTGAAATTAATGTGTGCCATGTCAGGGGTCTACGGTTTGTGGTAGTAACTTAATAC 3039
 Db 130 CTCCTGAAATTAATGTGTGCCATGTCAGGGGTCTACGGTTTGTGGTAGTAACTTAATAC 71
 QY 3040 CTTTACCCCAATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCAT 3099
 Db 70 CTTTACCCCAATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCAT 11
 QY 3100 AGAATTAAT 3109
 Db 10 AGAATTAAT 1

RESULT 133

DA456165

LOCUS

DEFINITION DA456165 CTONG3 Homo sapiens cDNA clone CTONG3001215 5', mRNA

ACCESSION DA456165

VERSION DA456165.1 GI:81187469

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

1 (bases 1 to 574)

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,

Yamashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaguri, H.,

Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,

Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuna, M.,

Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,

Tanabe, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.

Diversification and Characterization of Putative Alternative

Promoters of Human Genes

Genome Res. 16 (1), 55-65 (2006)

16344560

Contact: Takao Isogai

FLJ Project (HRI Team)

Helix Research Institute

2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology

Developmental Organization, Japan); cDNA library construction:

Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,

Research Association for Biotechnology (RAB) and Biotechnology

Center, National Institute of Technology and Evaluation; 3'-end one

pass sequencing: RAB.

JOURNAL PUBMED

COMMENT

FEATURES

source

ORIGIN

FEATURES
source
Location/Qualifiers
1. 574
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG3001215"
/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG3"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 13.8%; Score 430; DB 9; Length 574;
Best Local Similarity 99.8%; Pred. No. 1.5e-220; Indels 1; Gaps 1;
Matches 550; Conservative 0; Mismatches 0

QY 24 AGGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCCGCTTCCGCGCTGGGTGTTCACC 83
Db 24 AGGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCCGCTTCCGCGCTGGGTGTTCACC 83

QY 84 GCGCGCGCGCGCGAGAGCCACTGCAACACGAGCCGAGTGGAGCGCGCGAGCATGAA 143
Db 84 GCGCGCGCGCGCGAGAGCCACTGCAACACGAGCCGAGTGGAGCGCGCGAGCATGAA 143

QY 144 GC-GGCGAGCGCGCTCCATAGCGCAGTCCGGACGTCGCGCGCGCGCGCGGGAAG 202
Db 144 GCNGGCGCAGCGCGCTCCATAGCGCAGTCCGGACGTCGCGCGCGCGCGGGAAG 203

QY 203 GAAATGCAACATGGCAGCAGCAATGGAACACAGACAGTGGGTGTTCAGATATTGAAA 262
Db 204 GAAATGCAACATGGCAGCAGCAATGGAACACAGACAGTGGGTGTTCAGATATTGAAA 263

QY 263 CTGCGGACTCTGAGGAGAAATTGAATCAAGATCGGCTTAATTGAGCGCTTTTATG 322
Db 264 CTGCGGACTCTGAGGAGAAATTGAATCAAGATCGGCTTAATTGAGCGCTTTTATG 323

QY 323 TTGAGCGGTATCTTGAGTCACTTAAAGCTGCTTCCGATACCAAGAAATATCATG 382
Db 324 TTGAGCGGTATCTTGAGTCACTTAAAGCTGCTTCCGATACCAAGAAATATCATG 383

QY 383 GCTACATGATGGCTAAGGACACCATGATTTATGTTTGTGAAGAGGAATGATCCAGATG 442
Db 384 GCTACATGATGGCTAAGGACACCATGATTTATGTTTGTGAAGAGGAATGATCCAGATG 443

QY 443 GACCTCATTCAGACAGAAATCTATTACCTTGCCATGTCGTGAGAACAGAGAAATACAC 502
Db 444 GACCTCATTCAGACAGAAATCTATTACCTTGCCATGTCGTGAGAACAGAGAAATACAC 503

QY 503 TGTTTTTATCTGAAATCCCAAACTATCAATAGACGACGCTTAATGCTCTCTTGA 562
Db 504 TGTTTTTATCTGAAATCCCAAACTATCAATAGACGACGCTTAATGCTCTCTTGA 563

QY 563 AGCCTCTTTTG 573
Db 564 AGCCTCTTTTG 574

RESULT 134
DA649369
LOCUS
DEFINITION
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS
1 (bases 1 to 548)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,

Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Pujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T., and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Iobogal
PLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3966
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1. 548
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MESAN202496"
/cell_type="normal mesangial cells (NHMC56046-2)"
/clone_lib="MESAN2"
/note="Vector: pME18SFL3; primary culture, normal
mesangial cells"

ORIGIN
Query Match 13.6%; Score 424; DB 9; Length 548;
Best Local Similarity 99.8%; Pred. No. 2.6e-217;
Matches 544; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGCTGTACTGCGCGCTGCTTTCTTAGTCCG 60
Db 4 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGCTGTACTGCGCGCTGCTTTCTTAGTCCG 63

QY 61 GCTTCCGCGCTGGGTGTTCACCGCGCGCGCGAGGAGGCGCTTCCATAGCGCACGTCCG 120
Db 64 GCTTCCGCGCTGGGTGTTCACCGCGCGCGCGAGGAGGCGCTTCCATAGCGCACGTCCG 123

QY 121 GAGTGGAGCGCGCGCAGCATGAAAGCGCGCGAGGCGCGCTTCCATAGCGCACGTCCG 179
Db 124 GAGTGGAGCGCGCGCAGCATGAAAGCGCGCGAGGCGCGCTTCCATAGCGCACGTCCG 183

QY 180 GTCCGCGCGCGCGCGGAGGAAATGCAATGCAATGCGCAGCAGCAATGGAACAGAAC 239
Db 184 GTCCGCGCGCGCGCGGAGGAAATGCAATGCAATGCGCAGCAGCAATGGAACAGAAC 243

QY 240 GCTGGGTGTGAGATATTGAACTCGCGAGCTGTGAGGAGAAATTTGAATCAGAGATCG 299
Db 244 GCTGGGTGTGAGATATTGAACTCGCGAGCTGTGAGGAGAAATTTGAATCAGAGATCG 303

QY 300 GCCTAAATTTGAGCGCTTTTATGTTTGTGAGCGGTATTCTTGGAGTCAAGTCTTAAAGCTGCT 359
Db 304 GCCTAAATTTGAGCGCTTTTATGTTTGTGAGCGGTATTCTTGGAGTCAAGTCTTAAAGCTGCT 363

QY 360 TGCCGATACCAAGAAATATCATGATGCTTAAAGCGCAGCAGCAATGGAACAGAAC 419
Db 364 TGCCGATACCAAGAAATATCATGATGCTTAAAGCGCAGCAGCAATGGAACAGAAC 423

QY 420 TGTGAAGAGAAATGATCCAGATGCGCTTCAATCAGACAGAAATCTTATACCTTGGCATGTC 479
Db 424 TGTGAAGAGAAATGATCCAGATGCGCTTCAATCAGACAGAAATCTTATACCTTGGCATGTC 483

QY 480 TGGTGAGAACAGAGAAATATACACTGTTTATTTCTGAAATTCCTCAATATAGAGC 539
Db 484 TGGTGAGAACAGAGAAATATACACTGTTTATTTCTGAAATTCCTCAATATAGAGC 543

```

Qy 540 AGCAG 544
Db 544 AGCAG 548

RESULT 135
LOCUS DA392311
DEFINITION BRTHA2 Homo sapiens cDNA clone BRTHA2030255 5', mRNA
ACCSSION DA392311
VERSION DA392311.1 GI:80576008
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo
1 (bases 1 to 592)
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushiya, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction.
Helix Research Institute (HRI); 5'-end one pass sequencing. HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing. RAB.
Location/Qualifiers
1..592
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRTHA2030255"
/tissue_type="thalamus"
/clone_lib="BRTHA2"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 13.6%; Score 424; DB 9; Length 592;
Best Local Similarity 100.0%; Pred. No. 2.6e-217; Indels 0; Gaps 0;
Matches 424; Conservative 0; Mismatches 0;

Qy 171 GTCCGACGTCGCGGGCGGGCGGGGAGGAGGAAATGCAATGCGCAGCAATGGA 230
Db 169 GTCCGACGTCGCGGGCGGGCGGGGAGGAGGAAATGCAATGCGCAGCAATGGA 228

Qy 231 AACAGACAGCTGGGTGTGAGATATTGAACTGCGGACTGTGAGAGAAATATTGAATC 290
Db 229 AACAGACAGCTGGGTGTGAGATATTGAACTGCGGACTGTGAGAGAAATATTGAATC 288

Qy 291 ACAGCATCGGCTAAATGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAA 350
Db 289 ACAGCATCGGCTAAATGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAA 348

Qy 351 AAAGCTGCTCCGATACCCAGAAATATCATGGCTACATGGCTTAAGGCCACCATGGA 410
Db 349 AAAGCTGCTCCGATACCCAGAAATATCATGGCTACATGGCTTAAGGCCACCATGGA 408

us-10-825-632-2.olig.rst

Qy 411 TTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATTCAGACAGAATCTATTACCT 470
Db 409 TTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATTCAGACAGAATCTATTACCT 468

Qy 471 TGCATGTCCTGTTGAGAAACAGAGAAATACACTGTTTATTTCTGAAATTCGCAAACTAT 530
Db 469 TGCATGTCCTGTTGAGAAACAGAGAAATACACTGTTTATTTCTGAAATTCGCAAACTAT 528

Qy 531 CAATAGACAGCAGCTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTCAGGCAAC 590
Db 529 CAATAGACAGCAGCTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTCAGGCAAC 588

Qy 591 ACTG 594
Db 589 ACTG 592

RESULT 136
LOCUS BI462092
DEFINITION 857 bp mRNA linear EST 21-AUG-2001
603205272F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5271040 5',
mRNA sequence.
ACCSSION BI462092
VERSION BI462092.1 GI:15252748
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo
1 (bases 1 to 857)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11683 row: p column: 17
High quality sequence stop: 736.
Location/Qualifiers
1..857
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5271040"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site: BamHI; Site_2: SalI-XhoI
(gtcgag); Oligo: dm primed using primer
5'-TTTATTTTATTTTATTTT-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the cap-trapper method (Carninci, in
press, 2000). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 13.4%; Score 418; DB 2; Length 857;
Best Local Similarity 100.0%; Pred. No. 4.9e-214;
Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 201 AGCAAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTTGAGATATTGGA 260

```

(3.5-4.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D. "

ORIGIN	Query Match	Score	DB 5	Length
Best Local Similarity	13.3%	416	DB 5	474
Matches	99.8%	0	Mismatches	1
Conservative	0	Indels	0	Gaps
0				
QY	948	AGATGCCAGATCAGTGGAGTGGTACCTTTGTTCTCCAGAGAGAAATTTGATAGATATTC	1007	
Db	8	AGATGCCAGATCAGTGGAGTGGTACCTTTGTTCTCCAGAGAGAAATTTGATAGATATTC	67	
QY	1008	TGGCTATTGGTGGTGTCTCCAAAAGCTGAAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAAT	1067	
Db	68	TGGCTATTGGTGGTGTCTCCAAAAGCTGAAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAAT	127	
QY	1068	TCTATATGAAGAAAATGATGAATCTGAGGTGGAAAATTTATTCATGTTACATCCCTATGTT	1127	
Db	128	TCTATATGAAGAAAATGATGAATCTGAGGTGGAAAATTTATTCATGTTACATCCCTATGTT	187	
QY	1128	GGAACAGAGGGGAGGAGTTCATTCCTGTTATCTTAAACAGGTACAGCAAAATCTTAAAGT	1187	
Db	188	GGAACAGAGGGGAGGAGTTCATTCCTGTTATCTTAAACAGGTACAGCAAAATCTTAAAGT	247	
QY	1188	CACCTTTTAAAGATGTCAGAAAATTAATGATGATGCTGGAAGGAGGATCATAGATGTCATAGA	1247	
Db	248	CACCTTTTAAAGATGTCAGAAAATTAATGATGATGCTGGAAGGAGGATCATAGATGTCATAGA	307	
QY	1248	TAAGGAACCTAAATCAACCTTTTGAGATTCATTTTGAAGGAGTGAATATATGTCAGAGC	1307	
Db	308	TAAGGAACCTAAATCAACCTTTTGAGATTCATTTTGAAGGAGTGAATATATGTCAGAGC	367	
QY	1308	TGGATGACCTCTGAGGGGAAATATGCTTGGTGCATCTCTACTAGATCGTCCCAAGACTCG	1367	
Db	368	TGGATGACCTCTGAGGGGAAATATGCTTGGTGCATCTCTACTAGATCGTCCCAAGACTCG	427	
QY	1368	CCTACAGATAGTCTGATCTCACCTGAAATTTATTTATCCAGTGAAG	1414	
Db	428	CCTACAGATAGTCTGATCTCACCTGAAATTTATTTATCCAGTGAAG	474	

RESULT 138
DA297112 BRHP2 Homo sapiens cDNA clone BRHIP2010255 5', mRNA linear EST 03-NOV-2005
LOCUS DA297112 570 bp
DEFINITION DA297112.1 GI:79235254
ACCESSION DA297112
VERSION DA297112.1
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 570)

REFERENCE
AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986

TITLE
JOURNAL
PUBMED
COMMENT

Db	209	AGGAAATGCAATGCACGACCAATGAAACAGAACAGCTGGGTGTGAGATATTGA	268
QY	261	AACGCGGAGCTGTGAGGAGAAATTTGAATCAGAGATCGGCTTAATTTGGAGCCCTTTT	320
Db	269	AACGCGGAGCTGTGAGGAGAAATTTGAATCAGAGATCGGCTTAATTTGGAGCCCTTTT	328
QY	321	TGTTGAGGGTATTCTCGGAGTCACTTAAAGCTGTTGCGGATACCAAGAAATATCA	380
Db	329	TGTTGAGGGTATTCTCGGAGTCACTTAAAGCTGTTGCGGATACCAAGAAATATCA	388
QY	381	TGGCTACATGATGCTTAAGGCAACACATGATTTTCATGTTGTGAAGAGGATGATCCAGA	440
Db	389	TGGCTACATGATGCTTAAGGCAACACATGATTTTCATGTTGTGAAGAGGATGATCCAGA	448
QY	441	TGGACCTCATTCAGACAGATCTATTACCTTGCATGCTGTTGGAACAGAGAAATATAC	500
Db	449	TGGACCTCATTCAGACAGATCTATTACCTTGCATGCTGTTGGAACAGAGAAATATAC	508
QY	501	ACTGTTTATTTCTGAATTTCCAAAACCTATCAATAGACGACGAGTCTTAAATGCTCTTG	560
Db	509	ACTGTTTATTTCTGAATTTCCAAAACCTATCAATAGACGACGAGTCTTAAATGCTCTTG	568
QY	561	GAAGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAA	618
Db	569	GAAGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAA	626

RESULT 137
CF137069
LOCUS CF137069 474 bp mRNA linear EST 09-SEP-2003
DEFINITION UI-HF-BN0-amk-f-06-0-UI.r1 NIH MGC_50 Homo sapiens cDNA clone IMAGE:3090562 5', mRNA sequence.

ACCESSION CF137069
VERSION CF137069.1
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 474)

REFERENCE
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)

JOURNAL 8889548
PUBMED
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Louis Staudt
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.
Location/Qualifiers
1. 474
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3090562"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (LTI)"
/clone_lib="NIH MGC 50"
/note="Vector: p773D-pacI; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA

FEATURES
source

Email: flj-cdna@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction:
 Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing: RAB.

FEATURES
 source
 1..570
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRH2010255"
 /tissue="types=hippocampus"
 /clone_lib="BRH1P2"
 /note="Vector: pME18SFL3"

ORIGIN
 Query Match 13.2%; Score 413; DB 9; Length 570;
 Best Local Similarity 100.0%; Pred. No. 2.4e-211;
 Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 146 GGCACGCGCGCTCCATAGCGCACGTCGGGACGGTCCGGGCGGGCGGGGGAAGGAA 205
 Db 158 GGCACGCGCGCTCCATAGCGCACGTCGGGACGGTCCGGGCGGGCGGGGGAAGGAA 217
 QY 206 AATGCAACATGGCAGCAGCAATGGAACAGAAACAGACAGCTGGGTGTTGAGATATTTGAACTG 265
 Db 218 AATGCAACATGGCAGCAGCAATGGAACAGAAACAGACAGCTGGGTGTTGAGATATTTGAACTG 277
 QY 266 CGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCTAAATTCGGACCTTTTATGTTG 325
 Db 278 CGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCTAAATTCGGACCTTTTATGTTG 337
 QY 326 AGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGTTCGGATACAGAGCAATATCATGGCT 385
 Db 338 AGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGTTCGGATACAGAGCAATATCATGGCT 397
 QY 386 ACATGATGGCTAAGGCACACATGATTTTCATGTTGTGAGAGCAATGTCAGATGGAC 445
 Db 398 ACATGATGGCTAAGGCACACATGATTTTCATGTTGTGAGAGCAATGTCAGATGGAC 457
 QY 446 CTCATTACAGACAGATCTATTACCTTGCATGTCGTGTGAGACAGAGAAATACACTGT 505
 Db 458 CTCATTACAGACAGATCTATTACCTTGCATGTCGTGTGAGACAGAGAAATACACTGT 517
 QY 506 TTATTTCTGAATTCCTCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCT 558
 Db 518 TTATTTCTGAATTCCTCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCT 570

RESULT 139
 DA551227
 LOCUS
 DEFINITION
 DA551227 Homo sapiens cDNA clone HCHON2000347 5', mRNA
 sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1. (bases 1 to 530)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakisugi, H.,
 Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kuehida, N.,
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
 Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Diversification of Transcriptional Modulation: Large-scale
 Identification and Characterization of Putative Alternative

Promoters of Human Genes
 Genome Res. 16 (1), 55-65 (2006)
 16344560
 Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction:
 Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing: RAB.

FEATURES
 source
 1..530
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="HCHON200347"
 /cell_type="chondrocytes (HC)"
 /clone_lib="HCHON2"
 /note="Vector: pME18SFL3; primary culture, chondrocytes"

ORIGIN
 Query Match 13.1%; Score 408; DB 9; Length 530;
 Best Local Similarity 99.8%; Pred. No. 1.2e-208;
 Matches 528; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTCTACTCTCGCGCGCTCTTCTTAGTGCGG 60
 Db 2 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTCTACTCTCGCGCGCTCTTCTTAGTGCGG 61
 QY 61 CGTTGCGCGCTGGGTGTCACCGCGCGCGCGCGAGAGCCACTGCACACGAGGACCG 120
 Db 62 CGTTGCGCGCTGGGTGTCACCGCGCGCGCGCGAGAGCCACTGCACACGAGGACCG 121
 QY 121 GAGTGAGGCGCGCGCAGCATGAAGC-GGCGAGCGCGCGCTCCATAGCGCAGTCGGAGCG 179
 Db 122 GAGTGAGGCGCGCGCAGCATGAAGCNGCGCGCGCGCTCCATAGCGCAGTCGGAGCG 181
 QY 180 GTCGCGCGCGCGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAAACAGAACCA 239
 Db 182 GTCGCGCGCGCGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAAACAGAACCA 241
 QY 240 GCTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGAGAAATTTGAATCAGAGATCG 299
 Db 242 GCTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGAGAAATTTGAATCAGAGATCG 301
 QY 300 GCTTAAATTTGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCT 359
 Db 302 GCTTAAATTTGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCT 361
 QY 360 TCCCGATACAGAAATATCATGCTACATGCTTAAGGCGCACCACCATGATTTTCATGTT 419
 Db 362 TCCCGATACAGAAATATCATGCTACATGCTTAAGGCGCACCACCATGATTTTCATGTT 421
 QY 420 TGTGAAGGAAATGATCCAGATGGAACCTCATTCAGACAGAAATTTTACCTTCCCATGTC 479
 Db 422 TGTGAAGGAAATGATCCAGATGGAACCTCATTCAGACAGAAATTTTACCTTCCCATGTC 481
 QY 480 TGTGAGAACACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAAACT 528
 Db 482 TGTGAGAACACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAAACT 530

RESULT 140
 DA436183
 LOCUS
 DEFINITION
 DA436183 CTONG2 Homo sapiens cDNA clone CTONG2003086 5', mRNA
 sequence.

	Db	522	TTCCCAAAACTATCAATAGACGAGCGTCTTAATGCTCTCTTGGAAGC	569
DA436183	VERSION	DA436183.1	GI:80867593	
EST.	KEYWORDS	Homo sapiens (human)		
SOURCE	ORGANISM	Homo sapiens		
REFERENCE	AUTHORS	Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.		
TITLE		Diversification and Characterization of Putative Alternative Promoters of Human Genes		
JOURNAL	PUBLISHED	Genome Res. 16 (1), 55-65 (2006)		
COMMENT		Contact: Takao Isogai FLJ Project (HRI Team) Helix Research Institute 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: flj-cdna@nifty.com		
FEATURES	source	Location/Qualifiers		
		1..569		
		/organism="Homo sapiens"		
		/mol_type="mRNA"		
		/db_xref="taxon:9606"		
		/clone="CTONG2003086"		
		/tissue type="tongue, tumor tissue"		
		/clone_lib="CTONG2"		
		/note="Vector: pME18SFL3"		
ORIGIN				
Query Match		13.1%; Score 408; DB 9; Length 569;		
Best Local Similarity		100.0%; Pred. No. 1.2e-208;		
Matches		Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	158	CTCATAGCGCACGTGGGACCGTTCGGGCGGGCGGGGGAAGGAAATGCCAATGG	217	
Db	162	CTCATAGCGCACGTGGGACCGTTCGGGCGGGCGGGGGAAGGAAATGCCAATGG	221	
Qy	218	CAGCAGCAATGGAAAACAGAACAGCTGGGTTCAGATATTTCGAACTCGGCAGTGTGAGG	277	
Db	222	CAGCAGCAATGGAAAACAGAACAGCTGGGTTCAGATATTTCGAACTCGGCAGTGTGAGG	281	
Qy	278	AGAATATTGAATCAGAGGATCGGCTTAATTTGAGCGCTTTTATGTTGAGCGGTATTCCT	337	
Db	282	AGAATATTGAATCAGAGGATCGGCTTAATTTGAGCGCTTTTATGTTGAGCGGTATTCCT	341	
Qy	338	GGAGTCAGCTTAAAAAGCTGCTTCCGATACAGAAAATATCATGGCTACATGATGGCTA	397	
Db	342	GGAGTCAGCTTAAAAAGCTGCTTCCGATACAGAAAATATCATGGCTACATGATGGCTA	401	
Qy	398	AGGCACCACATGATTTCAATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACA	457	
Db	402	AGGCACCACATGATTTCAATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACA	461	
Qy	458	GAACTATTACCTTGCATGCTGTGGTGAAGACAGAGAAATACACTGTTTTATTCTGAAA	517	
Db	462	GAACTATTACCTTGCATGCTGTGGTGAAGACAGAGAAATACACTGTTTTATTCTGAAA	521	
Qy	518	TTCCCAAACTATCAATAGACGAGCTCTTAATGCTCTCTTGGAAGC	565	

```

QY 419 TTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTTATTACCTTGCCATGT 478
    |||
Db 427 TTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTTATTACCTTGCCATGT 486
QY 479 CTGGTGAGAACAGAGAAATACACTGTTTATTCTGAAATTTCCCAAACTATCAATAGAG 538
    |||
Db 487 CTGGTGAGAACAGAGAAATACACTGTTTATTCTGAAATTTCCCAAACTATCAATAGAG 546
QY 539 CAGCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTT 582
    |||
Db 547 CAGCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTT 590

RESULT 142
BI548490
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BI548490
603191203P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5262336 5',
mRNA sequence.
BI548490
EST.
BI548490.1 GI:15435802
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 835)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11661 row: f column: 01
High quality sequence stop: 795.
Location/Qualifiers
1. .835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5262336"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH MGC 95"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+) Site_1: BamHI; Site_2: SalI-XhoI
(gtccag); Oligo 4m primed using primer
5'-TTTATTTTATTTTATTTVN-3', size-selected for average
insert size 2.5 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this
is a NIH_MGC Library."

Query Match
Best Local Similarity 12.9%; Score 404; DB 2; Length 835;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 AGGAAATGCAACATGCACAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTGA 260
Db 229 AGGAAATGCAACATGCACAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTGA 288
QY 261 AACTGCCGACTGTGAGAGAAATTTGAATCAGAGATCGCCCTAAATTTGGAGCCTTTT 320
    |||

```

```

Db 289 AACTGCCGACTGTGAGAGAAATTTGAATCAGAGATCGGCCTAAATTTGGAGCCTTTT 348
QY 321 TGTGAGCGGTATTCTCGAGTCAAGCTTTAAAAGCTGCTTGCCGATACCAAGAAATATCA 380
    |||
Db 349 TGTGAGCGGTATTCTCGAGTCAAGCTTTAAAAGCTGCTTGCCGATACCAAGAAATATCA 408
    |||
QY 381 TGGCTACATGATGGCTAAGGCACACACATGATTTTCATGTTTGTGAGAGGAATGATCCAGA 440
    |||
Db 409 TGGCTACATGATGGCTAAGGCACACACATGATTTTCATGTTTGTGAGAGGAATGATCCAGA 468
    |||
QY 441 TGGACCTCATTTCAGACAGAACTTATACCTTGCCATGCTGCTGGTGGAGAAAGAAATAC 500
    |||
Db 469 TGGACCTCATTTCAGACAGAACTTATACCTTGCCATGCTGCTGGTGGAGAAAGAAATAC 528
    |||
QY 501 ACTGTTTTTATCTGAATTTCCCAAACTATCAATAGACAGCAGTCTTTAATGCTCTCTTG 560
    |||
Db 529 ACTGTTTTTATCTGAATTTCCCAAACTATCAATAGACAGCAGTCTTTAATGCTCTCTTG 588
    |||
QY 561 GAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTATGGAA 604
    |||
Db 589 GAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTATGGAA 632
    |||

```

```

RESULT 143
AA769522/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AA769522
n235c06.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1289774 3',
mRNA sequence.
AA769522
EST.
AA769522.1 GI:2820760
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 399)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1011 Std Error: 0.00
Seq primer: -40ml3 fwd Rr from Amersham
High quality sequence stop: 394.
Location/Qualifiers
1. 399
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1289774"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/clone_lib="NCI_CGAP GCB1"
/notes="Vector: p7T3D-Faci; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was prepared from human tonsillar cells
enriched for germinal center B cells by flow sorting
(CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA
synthesis was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGGCTCATTTTTTTTTTTT-3',
1. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I

```

FEATURES

source

Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

Location/Qualifiers
source
1. .599
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TRACH2015254"
/tissue_type="trachea"
/clone_lib="TRACH2"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 12.8%; Score 398; DB 9; Length 599;
Best Local Similarity 100.0%; Pred. No. 3.3e-203; Indels. 0; Gaps 0;
Matches 398; Conservative 0; Mismatches 0;
QY 170 CGTCGGGACGGTCCGGGGCGGGGGGGAAGGAAATGCAACATGGCAGCAGCAATGG 229
Db 167 CGTCGGGACGGTCCGGGGCGGGGGGGAAGGAAATGCAACATGGCAGCAGCAATGG 226
QY 230 AAACAGAAACAGCTGGGTGTTGAGATATTTCAAACTGCGGACTGTGAGGAGATATTTGAAT 289
Db 227 AAACAGAAACAGCTGGGTGTTGAGATATTTCAAACTGCGGACTGTGAGGAGATATTTGAAT 286
QY 290 CACAGATCGGCTTAAATTTGAGGCTTTTATGTTGAGCGGTATTTCTCGGAGTCAGCTTA 349
Db 287 CACAGATCGGCTTAAATTTGAGGCTTTTATGTTGAGCGGTATTTCTCGGAGTCAGCTTA 346
QY 350 AAAAGCTGTTGCGGATACCAAGAAATATCATGCTACATGATGGCTTAAGGACCAACATG 409
Db 347 AAAAGCTGTTGCGGATACCAAGAAATATCATGCTACATGATGGCTTAAGGACCAACATG 406
QY 410 ATTTCATGTTTGAAGAGGAAATGATCCAGATGGACCTCATTCAGACAGAAATTTATPACC 469
Db 407 ATTTCATGTTTGAAGAGGAAATGATCCAGATGGACCTCATTCAGACAGAAATTTATPACC 466
QY 470 TTGCCATGTCGTGGTGAAGAAATACACACTGTTTATTTCTGAAATTTCCCAAACTA 529
Db 467 TTGCCATGTCGTGGTGAAGAAATACACACTGTTTATTTCTGAAATTTCCCAAACTA 526
QY 530 TCAATAGACGACGCTTAAATGCTCTCTCTGGAAGCCT 567
Db 527 TCAATAGACGACGCTTAAATGCTCTCTCTGGAAGCCT 564

RESULT 145
CN427234 738 bp mRNA linear EST 16-MAY-2004
LOCUS 17000600021819 GRN_PRENUE Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN427234
ACCESSION CN427234.1 GI:47414828
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 738)
Branderberger, K., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebrowski, J and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
JOURNAL 15146197
PUBMED
COMMENT Contact: Branderberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760

and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 12.8%; Score 399; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 9.2e-204; Indels 0; Gaps 0;
Matches 399; Conservative 0; Mismatches 0;
QY 2720 GAAAGCCATATGTTTACAGATCTATCTCCTCAGGAGACACAGCATAAGAGTTCTCTGAAT 2779
Db 399 GAAAGCCATATGTTTACAGATCTATCTCCTCAGGAGACACAGCATAAGAGTTCTCTGAAT 340
QY 2780 CGGAGAAACATTATGAACCTGCACTTTTTCACCTTCAAGAAACCTTGATCAGTA 2839
Db 339 CGGAGAAACATTATGAACCTGCACTTTTTCACCTTCAAGAAACCTTGATCAGTA 280
QY 2840 TTGCTGCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGATATACACTGGCT 2899
Db 279 TTGCTGCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGATATACACTGGCT 220
QY 2900 ATTTAAACCAATGAGGAGGTTTAATCAACAGAAACACAGAAATTCATCATCACTTTTGA 2959
Db 219 ATTTAAACCAATGAGGAGGTTTAATCAACAGAAACACAGAAATTCATCATCACTTTTGA 160
QY 2960 TACCTGCCATGTAACTACTCTCTGAAATAAATGTGTGCCATGCGAGGGTCTACGGT 3019
Db 159 TACCTGCCATGTAACTACTCTCTGAAATAAATGTGTGCCATGCGAGGGTCTACGGT 100
QY 3020 TTGTGTGTAGTAATCTAATACCTTTAACCCCACTGCTCAAAATCAAAATGATATATTCCT 3079
Db 99 TTGTGTGTAGTAATCTAATACCTTTAACCCCACTGCTCAAAATCAAAATGATATATTCCT 40
QY 3080 GAGAGACCCAGCAATACATAAGAAATTAATAAAAAA 3118
Db 39 GAGAGACCCAGCAATACATAAGAAATTAATAAAAAA 1

RESULT 144
DB204093 599 bp mRNA linear EST 04-DEC-2005
LOCUS DB204093 TRACH2 Homo sapiens cDNA clone TRACH2015254 5', mRNA sequence.
DEFINITION DB204093
ACCESSION DB204093.1 GI:83245645
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 599)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi, Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
JOURNAL 16344560
PUBMED
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,

Email: rbrandenberger@geron.com
Insert Length: 738 Std Error: 0.00.
Location/Qualifiers
1. .738
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated HES cell line H7"
/clone_lib="GSI PRENEU"
/note="oligo at primed, full-length enriched cDNA library
from HES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."

FEATURES
source

ORIGIN

Query Match 12.7%; Score 395; DB 8; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.4e-201;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1938 GAAGATCCACACTGTGTGTCCTTTTAAAGCTATCAAGTCCTGAAGATGACCAACTTG 1997
Db 9 GAAGATCCACACTGTGTGTCCTTTTAAAGCTATCAAGTCCTGAAGATGACCAACTTG 68
QY 1998 CAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATAC 2057
Db 69 CAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATAC 128
QY 2058 TCCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTACATTTGTATGGATGCTCTA 2117
Db 129 TCCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTACATTTGTATGGATGCTCTA 188
QY 2118 CAAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGCTGTTTCATATATGGTGG 2177
Db 189 CAAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGCTGTTTCATATATGGTGG 248
QY 2178 TCCTCAGGTGAGTGTGTGAATATCGTTTAAAGAGTCAAGTATTTCCGTTGAATAC 2237
Db 249 TCCTCAGGTGAGTGTGTGAATATCGTTTAAAGAGTCAAGTATTTCCGTTGAATAC 308
QY 2238 CTTAGCCTCTCAGTGTATGCTGTGTAGTATAGACACAGGATCTGTACCGAGG 2297
Db 309 CTTAGCCTCTCAGTGTATGCTGTGTAGTATAGACACAGGATCTGTACCGAGG 368
QY 2298 GCTTAAATTTGAAGCGCCTTTAAATATATAATGG 2332
Db 369 GCTTAAATTTGAAGCGCCTTTAAATATATAATGG 403

RESULT 146

AA278625
LOCUS AA278625 493 bp mRNA linear EST 15-AUG-1997
DEFINITION z578911.t1 NCI CGAP GCBI Homo sapiens cDNA clone IMAGE:703652 5',
similar to SW:DPP4_RAT P14740 DIPEPTIDYL PEPTIDASE IV ;, mRNA
sequence.
AA278625
VERSION AA278625.1 GI:1919945
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 817 Std Error: 0.00

Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 493.
Location/Qualifiers
1. .493

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:703652"
/tissue_type="germlinal center B cell"
/lab_host="DH10B"
/clone_lib="NCI CGAP GCBI"
/note="Vector: pT73B-Pacl; Site 1: Not I; Site 2: Eco RI;
1st strand cDNA was prepared from human tonsillar cells
enriched for germinal center B cells by flow sorting
(CD20+, IGD-), provided by Dr. Louis M. Staudt (NCI), Dr.
David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA
synthesis was primed with a Not I oligo (GT TTTT-3',
5'-GGTTACCAATCTGAAGTGGGCGGCTCATTTTCTTTT-3',
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 12.6%; Score 393; DB 1; Length 493;
Best Local Similarity 100.0%; Pred. No. 1.7e-200;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2333 GTCAAATGAAATTCAGATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGATT 2392
Db 1 GTCAAATGAAATTCAGATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGATT 60
QY 2393 TCATTGACTTATAGATCGTGTGGGCATCCACGCGTGGTCTATGGAGGATACCTCTCCCTGA 2452
Db 61 TCATTGACTTATAGATCGTGTGGGCATCCACGCGTGGTCTATGGAGGATACCTCTCCCTGA 120
QY 2453 TGGCATTAAATGACAGAGTTCAGATATCTTCAGGGTGTCTATTTGTTGGGGCCCCAGTCACTC 2512
Db 121 TGGCATTAAATGACAGAGTTCAGATATCTTCAGGGTGTCTATTTGTTGGGGCCCCAGTCACTC 180
QY 2513 TGTGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCCCTGACCAAGT 2572
Db 181 TGTGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCCCTGACCAAGT 240
QY 2573 AACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGACCAA 2632
Db 241 AACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGACCAA 300
QY 2633 ATCGTTTACTGCTCTTACATGTTTCTCTGATGAGATGTCATTTTGCACATACAGTA 2692
Db 301 ATCGTTTACTGCTCTTACATGTTTCTCTGATGAGATGTCATTTTGCACATACAGTA 360
QY 2693 TATTACTGAGTTTTTTAGTGGGCTGGAAAGC 2725
Db 361 TATTACTGAGTTTTTTAGTGGGCTGGAAAGC 393

RESULT 147

AA312210/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT


```

QY 1072 TATGAAGAAATGATGAATCTGAGGTGGAATATTATTCATCTACATCCCTATGTTGGA 1131
DB 402 TATGAAGAAATGATGAATCTGAGGTGGAATATTATTCATCTACATCCCTATGTTGGA 461
QY 1132 ACAAGGAGGCGAGATTTCATTCCTGTTTAAACAGGTACAGCAATCTTAAGTCACT 1191
DB 462 ACAAGGAGGCGAGATTTCATTCCTGTTTAAACAGGTACAGCAATCTTAAGTCACT 521
QY 1192 TTTAAGATGTCAGAAATTAATGATGCTG 1222
DB 522 TTTAAGATGTCAGAAATTAATGATGCTG 552

RESULT 149
BG119202 798 bp mRNA linear EST 08-MAY-2001
LOCUS 602690257F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822550 5',
DEFINITION mRNA sequence.
ACCESSION BG119202
VERSION BG119202.1 GI:13998389
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 798)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10731 row: a column: 15
High quality sequence stop: 796.
FEATURES
    source
        1..798
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4822550"
            /lab_host="NH108 97"
            /clone_id="NIH_MGC_97"
            /note="Organ: testis; Vector: pBluescriptR (modified
            pUCtag); Site_1: BamHI; Site_2: SalI-XhoI
            S-1:TTTTTTTTTTTTTTVN-3", size-selected for average
            insert size 2.2 Kb and normalized to ROT 5. This is a
            primary library enriched for full-length clones and
            constructed using the Cap-trapper method (Carninci, in
            preparation). Library constructed by M. Brownstein
            (NIH/NHGRI, National Institutes of Health). Note: this is
            a NIH_MGC Library."

ORIGIN
Query Match 12.4%; Score 388; DB 2; Length 798;
Best Local Similarity 99.8%; Pred. No. 8.9e-198;
Matches 438; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 201 AGGAAATGCAACATGGCAGCAATGGAACAGACAGCTGGGTGTTGAGATATTGA 260
DB 359 AGGAAATGCAACATGGCAGCAATGGAACAGACAGCTGGGTGTTGAGATATTGA 418
QY 261 AACTCGGAGTGTGAGGAGAAATTTGAATCACAGGATCGGCTTAATTCGAGCCCTTTT 320
DB 419 AACTCGGAGTGTGAGGAGAAATTTGAATCACAGGATCGGCTTAATTCGAGCCCTTTT 478

```

```

QY 321 TGTGAGCGGTATTCTCGAGTCAGCTTAAAAAGCTGCTTCCCGATACCGAATATCA 380
DB 479 TGTGAGCGGTATTCTCGAGTCAGCTTAAAAAGCTGCTTCCCGATACCGAATATCA 538
QY 381 TGGCTACATGATGGCTAAGGACCAACATGATTTTCATGTTTGTGAAGGAGGATGATCCAGA 440
DB 539 TGGCTACATGATGGCTAAGGACCAACATGATTTTCATGTTTGTGAAGGAGGATGATCCAGA 598
QY 441 TGGACCTCATTCCAGACAGAACTATTACCTTCCCATCTATGCTGCTGAGAACAGAAAATAC 500
DB 599 TGGACCTCATTCCAGACAGAACTATTACCTTCCCATCTATGCTGCTGAGAACAGAAAATAC 658
QY 501 ACTGTTTATTCTGAAATTCGCAAACTATCATATAGAGCAGCAGTCTTAATGCTCTCTTG 560
DB 659 ACTGTTTATTCTGAAATTCGCAAACTATCATATAGAGCAGCAGTCTTAATGCTCTCTTG 718
QY 561 GAAGCCTCTTTTGGATCTTTTTCAGGCAACATGAGACTATGGAATGATTTCTCGAGAAGA 620
DB 719 GAAGCCTCTTTTGGATCTTTTTCAGGCAACATGAGACTATGGAATGATTTCTCGAGAAGA 778
QY 621 AGAATCTATTAGAGAAAGA 639
DB 779 AGAATCTATTAGAGAAAGA 797

RESULT 150
DA184650 555 bp mRNA linear EST 03-NOV-2005
LOCUS DA184650 BRAMY3 Homo sapiens cDNA clone BRAMY3003323 5', mRNA
DEFINITION sequence.
ACCESSION DA184650
VERSION DA184650.1 GI:79178555
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 555)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Teuricani, K., Wakisugi, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Megatsuma, M.,
Murakawa, K., Iehida, S., Ishibashi, T., Takahashi, R., Tanase, T.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, J., A.,
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (FRI Team)
Helix Research Institute
2-6-7 Kazuoka-cho, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj_cdn@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation, 3'-end one
pass sequencing: RAB.
FEATURES
    Location/Qualifiers
        1..555
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="BRAMY3003323"
            /issue_type="amygdala"
            /clone_lib="BRAMY3"
            /note="Vector: pME18SFL3"

```

Thu Jun 22 09:04:26 2006

ORIGIN

Query Match 12.3%; Score 384; DB 9; Length 555;
Best Local Similarity 99.8%; Fred. No. 1.3e-195; Indels 0; Gaps 0;
Matches 434; Conservative 0; Mismatches 1;

Qy	106	CTGCAACACAGGACCGGAGTGGAGCGCGCAGCATGAAGCGGCGCAGGCGCGCTCCATAG	165
Db	102	CTGCAACACAGGACCGGAGTGGAGCGCGCAGCATGAAGCGGCGCAGGCGCGCTCCATAG	161
Qy	166	CGCACGTCCGGACCGTCCGGCGCGGCGCGGCGGGAAGAAATGCAACATGGCAGCAGCA	225
Db	162	CGCACGTCCGGACCGTCCGGCGCGGCGCGGCGGGAAGAAATGCAACATGGCAGCAGCA	221
Qy	226	ATGGAACACAGACAGCTGGGTCTTGAGATATTTGAAACTGCGGACTGTGAGGAGAATATT	285
Db	222	ATGGAACACAGACAGCTGGGTCTTGAGATATTTGAAACTGCGGACTGTGAGGAGAATATT	281
Qy	286	GAATCACAGGATCGGCCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAG	345
Db	282	GAATCACAGGATCGGCCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAG	341
Qy	346	CTTAAAAAGCTGTTCCCGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCACCA	405
Db	342	CTTAAAAAGCTGTTCCCGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCACCA	401
Qy	406	CATGATTTCAATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTAT	465
Db	402	CATGATTTCAATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCANACAGAATCTAT	461
Qy	466	TACCTTGCCCATGCTCGTGAGAACAGAGAAATACACTGTTTATTCTGAAATTCGCCAA	525
Db	462	TACCTTGCCCATGCTCGTGAGAACAGAGAAATACACTGTTTATTCTGAAATTCGCCAA	521
Qy	526	ACTATCAATAGAGCA	540
Db	522	ACTATCAATAGAGCA	536

Search completed: June 22, 2006, 08:05:45
Job time : 14083 secs

THIS PAGE BLANK (USPTO)

Thu Jun 22 09:04:25 2006

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2006, 03:55:04 ; Search time 17271 Seconds
(without alignments)
11552.079 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgtaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 6366136 seqs, 31973710525 residues

Word size : 1
Total number of hits satisfying chosen parameters: 12730834

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 150 summaries

Database : GenEmbl.*

- 1: gb env.*
- 2: gb pat.*
- 3: gb pi.*
- 4: gb pi.*
- 5: gb pr.*
- 6: gb ro.*
- 7: gb sts.*
- 8: gb sv.*
- 9: gb un.*
- 10: gb vi.*
- 11: gb ov.*
- 12: gb htg.*
- 13: gb in.*
- 14: gb om.*
- 15: gb ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3120	100.0	3120	2	AR651452 Sequence
2	3120	100.0	3127	5	AF221634 Homo sapi
3	3069	98.4	3127	2	CS196102 Sequence
4	3003	96.2	3106	2	AX342633 Sequence
5	2957	94.8	3143	2	AX354793 Sequence
6	2889	92.6	4829	2	AR631283 Sequence
7	2889	92.6	4829	2	AX608735 Sequence
8	2756	88.3	3030	5	AX354202 Homo sapi
9	2547	81.6	2649	2	AX354795 Sequence
10	2495	80.0	2671	2	AX172659 Homo sapi
11	2495	80.0	2671	2	AR631278 Sequence
12	2495	80.0	2671	2	AX608725 Sequence
13	2315	74.2	2797	2	AR448400 Sequence
14	2128	68.2	4523	2	AR631281 Sequence
15	2128	68.2	4523	2	AX608731 Sequence
16	2128	68.2	4676	2	AR631287 Sequence
17	2128	68.2	4676	2	AX608743 Sequence
18	2116	67.8	2842	2	AX405770 Sequence

19	2081	66.7	3125	5	BC030688
20	1984	63.6	4685	2	AR631288
21	1984	63.6	4685	2	AR631288
22	1963	62.9	2778	5	AK000290
23	1450	46.5	2161	2	BD157001
24	1450	46.5	2161	2	AX878058
25	1450	46.5	2161	5	AK027826
26	1377	44.1	2668	2	AX405771
27	1191	38.2	4309	2	AR631284
28	1191	38.2	4309	2	AR631284
29	1164	37.3	2510	2	AX338497
30	1023	32.8	1669	5	AR651454
31	1023	32.8	1669	5	AF221636
32	878	28.1	2098	2	CO721272
33	877	28.1	1356	2	AR631282
34	877	28.1	1356	2	AR631282
35	790	25.3	1197	5	AR651453
36	790	25.3	1197	5	AF221635
37	789	25.3	1083	2	AR651455
38	789	25.3	1083	5	AF221637
39	760	24.4	832	2	AR631286
40	760	24.4	832	2	AR631286
41	739	23.7	873	2	AX08741
42	708	22.7	735	2	AX524942
43	700	22.4	753	2	AX524935
44	672	21.5	1002	2	AX524939
45	501	16.1	631	2	BD149852
46	501	16.1	631	2	AX869790
47	472	15.1	620	2	AR631285
48	472	15.1	620	2	AX608739
49	383	12.3	198295	5	AC015129
50	339	10.9	147760	5	AC011846
51	209	9.9	823	2	AR263393
52	205	6.6	587	2	BD154822
53	205	6.6	587	2	AX874760
54	188	6.0	308	2	AX970984
55	188	6.0	308	2	BD109703
56	188	6.0	308	2	AR414150
57	167	5.4	168	2	AX522584
58	156	5.0	176282	12	AC068507
59	82	2.6	89	2	BD036103
60	82	2.6	89	2	AR736364
61	82	2.6	89	2	AX900570
62	82	2.6	477	7	BV448370
63	81	2.6	222034	12	AC157219
64	81	2.0	224227	12	AC106509
65	56	1.8	4799	6	BC043124
66	56	1.8	4799	6	BC059222
67	54	1.7	305843	12	AC144143
68	48	1.5	305843	12	AC144143
69	47	1.5	186144	6	AC161443
70	46	1.5	60	2	CQ545450
71	44	1.4	197073	12	AC171824
72	42	1.3	116580	6	AC112162
73	42	1.3	155733	6	AC161366
74	42	1.3	162528	6	AC122463
75	34	1.1	612	2	AR372392
76	33	1.1	72149	12	AC122776
77	32	1.0	1449	11	CR353491
78	29	0.9	194188	6	AC163387
79	26	0.8	2200	11	BC108576
80	25	0.8	25	2	CS196106
81	25	0.8	51	2	CQ002331
82	24	0.8	9154	2	CQ572930
83	24	0.8	40388	2	CQ608996
84	24	0.8	79011	12	AC015386
85	24	0.8	102653	4	AC090436
86	24	0.8	127686	4	AC090434
87	24	0.8	172372	13	AC010110
88	24	0.8	298235	12	AC127924
89	24	0.8	301691	13	AE003475
90	23	0.7	23	2	CS196104
91	23	0.7	23	2	AR631300

92	23	0.7	29	2	AX608768	Sequence	Unclassified.
93	23	0.7	632	7	BV055847	S212P6100	1 (bases 1 to 3120)
94	23	0.7	635	7	BV311981	S236P628F	Abbott, C.A. and Gorrell, M.D.
C 95	23	0.7	1275	2	CS211305	AB056577	Deptideyl peptidases
C 96	23	0.7	1275	6	AB056577	Measurice	Patent: US 6881564-A 2 19-APR-2005;
C 97	23	0.7	74579	11	AC149896	AC149892	The University of Sydney; Sydney;
C 98	23	0.7	98359	12	AC149892	Xenopus t	AUX;
99	23	0.7	213261	11	AC161463	Xenopus t	
100	22	0.7	22	2	AR631303	Gallus ga	Location/Qualifiers
C 101	22	0.7	22	2	AR631303	Sequence	1. .3120
C 102	22	0.7	22	2	AR631304	Sequence	/organism="unknown"
C 103	22	0.7	22	2	AX608771	Sequence	/mol_type="genomic DNA"
C 104	22	0.7	22	2	AX608772	Sequence	
C 105	22	0.7	39	2	AR631301	Sequence	
C 106	22	0.7	39	2	AX608769	Sequence	
107	22	0.7	2250	11	CT030588	Xenopus t	Query Match
108	22	0.7	3322	2	BD103383	Novel TTV	100.0%; Score 3120; DB 2; Length 3120;
109	22	0.7	3322	2	BD103384	Novel TTV	Best Local Similarity 100.0%; Pred. No. 0;
110	22	0.7	3322	10	AB049608	TT virus	Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
C 111	22	0.7	43049	5	HGDJ87818	Human DNA	
C 112	22	0.7	104012	12	AP000861	Homo sapi	
C 113	22	0.7	105350	12	AP000661	Homo sapi	
C 114	22	0.7	110000	4	AP000661	Continuation (41 o	
115	22	0.7	117345	5	AP001775	Homo sapi	
116	22	0.7	117445	12	AL160263	Homo sapi	
117	22	0.7	145028	12	CR8479173	Danio rer	
C 118	22	0.7	162862	12	CR847930	Zebrafish	
C 119	22	0.7	165540	12	AC158877	Bos tauru	
C 120	22	0.7	170051	4	AP005772	Oryza sat	
C 121	22	0.7	180005	12	AC136070	Rattus no	
C 122	22	0.7	187676	12	AC174595	Bos tauru	
C 123	22	0.7	195505	11	BX537287	Zebrafish	
C 124	22	0.7	196928	12	AP000847	Homo sapi	
C 125	22	0.7	207817	6	AC123934	Mus muscu	
C 126	22	0.7	210651	12	AC151257	Bos tauru	
C 127	22	0.7	210788	12	AC167594	Bos tauru	
128	22	0.7	214864	12	AC162627	Bos tauru	
C 129	22	0.7	222380	12	AC156845	Bos tauru	
C 130	22	0.7	229816	6	AC105304	Mus muscu	
131	22	0.7	231295	12	AC097286	Rattus no	
132	22	0.7	258530	12	CT030007	Danio rer	
133	21	0.7	294090	12	AC120770	Rattus no	
134	21	0.7	400	4	CQ690820	Sequence	
135	21	0.7	535	2	AR7865196	Cocos nuc	
136	21	0.7	535	2	AR280920	Sequence	
137	21	0.7	535	2	AR283416	Sequence	
138	21	0.7	535	2	AR344184	Sequence	
139	21	0.7	535	2	AR351385	Sequence	
140	21	0.7	535	2	AR453965	Sequence	
141	21	0.7	535	2	AR561553	Sequence	
142	21	0.7	535	2	AR632008	Sequence	
143	21	0.7	535	2	AR760254	Sequence	
144	21	0.7	535	2	AR773397	Sequence	
C 145	21	0.7	535	7	AV343295	Sequence	
C 146	21	0.7	802	2	CQ431300	Sequence	
C 147	21	0.7	855	2	CQ422442	Sequence	
148	21	0.7	1143	5	AF007822	Homo sapi	
149	21	0.7	1615	5	AL832166	Homo sapi	
150	21	0.7	2255	2	CS141101	Sequence	

RESULT 1	AR651452	3120 bp	DNA	linear	PAT 20-APR-2005
LOCUS	AR651452				
DEFINITION	Sequence	2 from patent US 6881564.			
ACCESSION	AR651452				
VERSION	AR651452.1	GI:62795938			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				

Qy 841 GACTGAAATGCTTTTATATACATAGCAACGATATTTTGGATATCTAAATCGTAAACAGAGAA 900
Db |||||
841 GACTGAAATGCTTTTATATACATAGCAACGATATTTTGGATATCTAAATCGTAAACAGAGAA 900
Qy |||||
901 GAAAGGAGACTCACTTATGTGACAAATGAGCTAGCCACATGGAAGAGATGCCAGATCA 960
Db |||||
901 GAAAGGAGACTCACTTATGTGACAAATGAGCTAGCCACATGGAAGAGATGCCAGATCA 960
Qy |||||
961 GCTGGAGTCCGTACCTTTGTTCTCCAAAGAAGAAATTTGATAGATATCTGGCTATTTGGTGG 1020
Db |||||
961 GCTGGAGTCCGTACCTTTGTTCTCCAAAGAAGAAATTTGATAGATATCTGGCTATTTGGTGG 1020
Qy |||||
1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Db |||||
1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Qy |||||
1081 AATGATGAATCTGAGGTGGAATTAATTCATGCTTACATCCCTATGTTTGGAAACCAAGGAGG 1140
Db |||||
1081 AATGATGAATCTGAGGTGGAATTAATTCATGCTTACATCCCTATGTTTGGAAACCAAGGAGG 1140
Qy |||||
1141 GCAGATTCATCCGTTTATCTTAAACACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATG 1200
Db |||||
1141 GCAGATTCATCCGTTTATCTTAAACACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATG 1200
Qy |||||
1201 TCAGAAATTAATGATGCTGGAAGGAGATCATAGATGTCATAGATAAGGAACCTAATTT 1260
Db |||||
1201 TCAGAAATTAATGATGCTGGAAGGAGATCATAGATGTCATAGATAAGGAACCTAATTT 1260
Qy |||||
1261 CAACTTTTGAAGATCTATTTTGAAGGAGTTCAATATATTTGCAGAGCTGGATGCACTCCT 1320
Db |||||
1261 CAACTTTTGAAGATCTATTTTGAAGGAGTTCAATATATTTGCAGAGCTGGATGCACTCCT 1320
Qy |||||
1321 GAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTCCACAGATAGATG 1380
Db |||||
1321 GAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTCCACAGATAGATG 1380
Qy |||||
1381 TTGATCTCACCTGAAATTAATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGACTC 1440
Db |||||
1381 TTGATCTCACCTGAAATTAATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGACTC 1440
Qy |||||
1441 ATTGAGTCAAGTCCCTGATCTGTGACGCCACTAATATCTATGAAGAAACAAACAGACATC 1500
Db |||||
1441 ATTGAGTCAAGTCCCTGATCTGTGAGGCCACTAATATCTATGAAGAAACAAACAGACATC 1500
Qy |||||
1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTTCCCAAAAGTCAGAGGAAATTTGAG 1560
Db |||||
1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTTCCCAAAAGTCAGAGGAAATTTGAG 1560
Qy |||||
1561 TTTATTTTTTGGCTCTGAAATGCAAAACAGGTTTTCCGTCAATTTATACAAAATTACATCTATT 1620
Db |||||
1561 TTTATTTTTTGGCTCTGAAATGCAAAACAGGTTTTCCGTCAATTTATACAAAATTACATCTATT 1620
Qy |||||
1621 TTTAAAGGAAGCAAAATATAAACGATCCAGTGGTGGGCTGCTGCTCCAAAGTGAATTTCAAG 1680
Db |||||
1621 TTTAAAGGAAGCAAAATATAAACGATCCAGTGGTGGGCTGCTGCTCCAAAGTGAATTTCAAG 1680
Qy |||||
1681 TGTCTCTATCAAAAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Db |||||
1681 TGTCTCTATCAAAAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Qy |||||
1741 GGATCTAATATATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGAC 1800
Db |||||
1741 GGATCTAATATATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGAC 1800
Qy |||||
1801 TCCCTTTTAGAGCATCACTGTACGTAAGTCAAGTTTAAATCTCTGAGAGGTTGACAAGG 1860
Db |||||
1801 TCCCTTTTAGAGCATCACTGTACGTAAGTCAAGTTTAAATCTCTGAGAGGTTGACAAGG 1860
Qy |||||
1861 CTGACTGACCGGGCTACTCACATTTCTGCTGCAATCAGTCAGCACTGTGACTCTTTTATA 1920
Db |||||
1861 CTGACTGACCGGGCTACTCACATTTCTGCTGCAATCAGTCAGCACTGTGACTCTTTTATA 1920
Qy |||||
1921 AGTAAAGTATAGTAAACCAAGAAGATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCT 1980

Db |||||
1921 AGTAAAGTATAGTAAACCAAGAAGATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCT 1980
Qy |||||
1981 GAAGATGACCCCAACTTTGCAAAACAAAGGAATTTTGGGCCACCATTTTTCGATTTCAAGCAGGT 2040
Db |||||
1981 GAAGATGACCCCAACTTTGCAAAACAAAGGAATTTTGGGCCACCATTTTTCGATTTCAAGCAGGT 2040
Qy |||||
2041 CCTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTGAAGTACTACTCTGGATTTTACA 2100
Db |||||
2041 CCTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTGAAGTACTACTCTGGATTTTACA 2100
Qy |||||
2101 TTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTCGGAAGAAATATCTCTACTGTG 2160
Db |||||
2101 TTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTCGGAAGAAATATCTCTACTGTG 2160
Qy |||||
2161 CTGTTTCATATATATGGTGGTCTCAGGTGCAAGTGGTGAATTAATCGGTTTTAAAGGAGTCAAG 2220
Db |||||
2161 CTGTTTCATATATGGTGGTCTCAGGTGCAAGTGGTGAATTAATCGGTTTTAAAGGAGTCAAG 2220
Qy |||||
2221 TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGTTAGTAGTAGTACAGCAAGG 2280
Db |||||
2221 TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGTTAGTAGTAGTACAGCAAGG 2280
Qy |||||
2281 GGATCCTGTCAACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATCGGTCAAATA 2340
Db |||||
2281 GGATCCTGTCAACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATCGGTCAAATA 2340
Qy |||||
2341 GAAATTCAGATCAGGTGGAAGGACTCCAAATATCTAGCTTTCTCGATATGATTTTCATTGAC 2400
Db |||||
2341 GAAATTCAGATCAGGTGGAAGGACTCCAAATATCTAGCTTTCTCGATATGATTTTCATTGAC 2400
Qy |||||
2401 TTAGATCGTGTGGGCATTCACCGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTA 2460
Db |||||
2401 TTAGATCGTGTGGGCATTCACCGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTA 2460
Qy |||||
2461 ATGCAGAGGTTCAGATATCTTTCAGGTTGCTATTTGCTGGGGCCCAAGTCACTCTGTGGATC 2520
Db |||||
2461 ATGCAGAGGTTCAGATATCTTTCAGGTTGCTATTTGCTGGGGCCCAAGTCACTCTGTGGATC 2520
Qy |||||
2521 TTTCTATGATACAGGATACACGGAACGTTTATATGCGTCACTCCCTGACAGAAATGAACAGGC 2580
Db |||||
2521 TTTCTATGATACAGGATACACGGAACGTTTATATGCGTCACTCCCTGACAGAAATGAACAGGC 2580
Qy |||||
2581 TATTTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGACCAAAATCGTTTA 2640
Db |||||
2581 TATTTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGACCAAAATCGTTTA 2640
Qy |||||
2641 CTGCTCTTACATGATGTTTCTCGATGAGATGTCCTATTTTGACATATCCAGTATATTAATCTG 2700
Db |||||
2641 CTGCTCTTACATGATGTTTCTCGATGAGATGTCCTATTTTGACATATCCAGTATATTAATCTG 2700
Qy |||||
2701 AGTTTTTTTAGTGAAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGAGAGACAC 2760
Db |||||
2701 AGTTTTTTTAGTGAAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGAGAGACAC 2760
Qy |||||
2761 AGCATTAAGTTCCTGAAATCGGGAGAACATTTAGAACTGCACTTTTTCGATCTACTCTCAA 2820
Db |||||
2761 AGCATTAAGTTCCTGAAATCGGGAGAACATTTAGAACTGCACTTTTTCGATCTACTCTCAA 2820
Qy |||||
2821 GAAAACTTTGGATCAGCTATTTGCTGCTCTTAAAGTGAATATAATTTTGAACCTGTGTAGAAC 2880
Db |||||
2821 GAAAACTTTGGATCAGCTATTTGCTGCTCTTAAAGTGAATATAATTTTGAACCTGTGTAGAAC 2880
Qy |||||
2881 TCTCTGGTATACACTGGCTATTTTAAACCAATGAGAGGTTTAAATCAACAGAAAAACACAGA 2940
Db |||||
2881 TCTCTGGTATACACTGGCTATTTTAAACCAATGAGAGGTTTAAATCAACAGAAAAACACAGA 2940
Qy |||||
2941 ATTGATCATCACATTTTGTATCTGCTCATGTAAACATCTACTCTCTGAAAAATAAATGTGGTG 3000
Db |||||
2941 ATTGATCATCACATTTTGTATCTGCTCATGTAAACATCTACTCTCTGAAAAATAAATGTGGTG 3000
Qy |||||
3001 CCATGAGGGGCTTACGGTTTGTGGTAGTAAATCTTAATACCTTAAACCCCAATGTCTCAAAA 3060
Db |||||

Db 3001 CCATGCGGGCTTACGGTTTGTGGTAGTAATCTAATACCTTAACCCCAATGCTCAAAA 3060
 Qy 3061 TCAAAATGATACATATTCCTGAGAGACCCAGCAATACCATAGATTAATTAATAAAAAAAA 3120
 Db 3061 TCAAAATGATACATATTCCTGAGAGACCCAGCAATACCATAGATTAATTAATAAAAAAAA 3120

RESULT 2
 AF221634 3127 bp mRNA linear PRI 05-NOV-2000
 LOCUS Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, complete cds.
 DEFINITION AF221634.1 GI:11095187
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 TITLE 1 (bases 1 to 3127)
 JOURNAL Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W. and Gorrell,M.D.
 PUBMED Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homologue, DPP8
 AUTHORS Eur. J. Biochem. 267 (20), 6140-6150 (2000)
 JOURNAL 11012666
 TITLE 2 (bases 1 to 3127)
 JOURNAL Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.
 PUBMED Direct Submission
 AUTHORS Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver Centre, Centenary Institute of Cell Biology and Cancer Medicine, Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia
 JOURNAL Location/Qualifiers

FEATURES
 source
 1..3127
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15q22"
 1..3127
 /gene="DPP8"
 214..2862
 /gene="DPP8"
 /note="serine protease; member of prolyl oligopeptidase S9b family belonging to enzyme clan SC; alternatively spliced forms of this mRNA are found in AF221635, AF221636 and AF221637"
 /codon_start=1
 /product="dipeptidyl peptidase 8"
 /protein_id="AAC29786.1"
 /db_xref="GI:11095188"
 /translation="MAAMETELQGVIEIFETADCEINIESQDRPKLEFFYVERYSWSQ LKLLADRIKHGMKAPDFMFKVRNDPDGPHSDRIYILAMSGENRENTLYSEI PKTINRAAVLMSKELLDFQATLDYGMYSREELLRRKRIKGTGVASDYHQSG TFLPQAGSLYHVDGPGQFTQOPLRNVLVETSCFNIRMDPKLCPADPDWIAF IHSN DIMISNITREERLTYVHNELANMEEDARSAGVATFVLQEFDRYSGYWCPKAETT PSGRILIRLYEENDESEVEI IHVTSFMLETRADSFRPKGTANPKVTFKQSEIM DAEGRIIDVDEKLIQPFILFEGVEYIARAGWTEPKYAWSIILDRSQRLQIVLIS IFAPECTVDMMERQRLIESVDSVTPLLIYIETTDIWINIHDFHVPQSHHEEIEF IFASECKTFRHLKYKITSILKESKYRSGSLPAPDFKPIKEBIAITSGEWEVLGR HGNSIQVDEIRLVLYFEKTKDPLHLLYVSVYVNPGEVRLTRDRGYSHCISQCHD FFTSKYNQPHCVKLSYLSKSPEDDPTCKTEFWATILDSAGPLPDYTPPEIFSFES TTGTLTGMLYKPHDLPQKKYPTVLFIYGGPQVQVNNRFRKVKYFRLNLTASLGVY VVVDDNNGSCHRLGFEKPGKYMKGQIEIDQVEGLQYLSRVDYFIDLRVHIGWSY GGYLSMALMQRSDIFRVAIGAPVTLWIFYDTGYTERYMGHPQDNQGYVLGVSVMQ AEKFPSPENRLLHLLHGFLENVHFAHTSILLISFLVRAGKPYDLQIYQPERHSIRVPES GEHYELHLLHYLQENLGSRIAAKVI"

ORIGIN

Query Match 100.0%; Score 3120; DB 5; Length 3127;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGTGTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCCGCTGCTTCTTAGTCCG 60
 Db 1 AAGTGTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCCGCTGCTTCTTAGTCCG 60
 Qy 61 CGTTCGCCCTCGGTGTGTACCCGCGCGCGCGCGAGGAAGCCACTGCAACACGAGACCG 120
 Db 61 CGTTCGCCCTCGGTGTGTACCCGCGCGCGCGCGAGGAAGCCACTGCAACACGAGACCG 120
 Qy 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGTTCGGGACGG 180
 Db 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGTTCGGGACGG 180
 Qy 181 TCCGGCGCGCGCGCGGGAAGAAATGCAACATGCGCAGCAGCAATGGAAACAGAACAG 240
 Db 181 TCCGGCGCGCGCGCGGGAAGAAATGCAACATGCGCAGCAGCAATGGAAACAGAACAG 240
 Qy 241 CTGGGTGTGAGATATTTGAAACTGCGGACTGTCAGGAGAAATATTGAATCACAGGATCG 300
 Db 241 CTGGGTGTGAGATATTTGAAACTGCGGACTGTCAGGAGAAATATTGAATCACAGGATCG 300
 Qy 301 CCTAAATTCGAGCCCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAAGCTGCTT 360
 Db 301 CCTAAATTCGAGCCCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAAGCTGCTT 360
 Qy 361 GCCGATACCAGAAAATATCATGCGCTACATGATGCTTAAGCCACCAATGATGATGAT 420
 Db 361 GCCGATACCAGAAAATATCATGCGCTACATGATGCTTAAGCCACCAATGATGATGAT 420
 Qy 421 GTGAAGAGGATGATCCAGATGCGACCTCATTCAGACAGAAATCTTACCTTGGCATGTCT 480
 Db 421 GTGAAGAGGATGATCCAGATGCGACCTCATTCAGACAGAAATCTTACCTTGGCATGTCT 480
 Qy 481 GGTGAGAACAGAAAATATCACTGTTTATTCGAAATCCCAAAATCAATCAATAGACA 540
 Db 481 GGTGAGAACAGAAAATATCACTGTTTATTCGAAATCCCAAAATCAATCAATAGACA 540
 Qy 541 GCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
 Db 541 GCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
 Qy 601 GGAATGTAATTCGAGAGAAAGAACTATTAAAGAGAAAGAAACGCATTGGAAACAGTCGA 660
 Db 601 GGAATGTAATTCGAGAGAAAGAACTATTAAAGAGAAAGAAACGCATTGGAAACAGTCGA 660
 Qy 661 ATTGCTTCTTACGATTATCAACAAAGGAAGTGGAAACATTTCTGTTTCAAGCCGCTAGTGA 720
 Db 661 ATTGCTTCTTACGATTATCAACAAAGGAAGTGGAAACATTTCTGTTTCAAGCCGCTAGTGA 720
 Qy 721 ATTTATCACTGTAAGAGATGGAGGCCCAAGAGGATTTACGCAACAACTTTAAAGGCCCAAT 780
 Db 721 ATTTATCACTGTAAGAGATGGAGGCCCAAGAGGATTTACGCAACAACTTTAAAGGCCCAAT 780
 Qy 781 CTAGTGGAAACTAGTGTGTCACCATACGCGATGGATCCAAATTTATGCCCCGCTGATCCA 840
 Db 781 CTAGTGGAAACTAGTGTGTCACCATACGCGATGGATCCAAATTTATGCCCCGCTGATCCA 840
 Qy 841 GACTGGATTGCTTTTATACATAGCAAGATTTTGGATATCTAAACATCGTAACCCAGANA 900
 Db 841 GACTGGATTGCTTTTATACATAGCAAGATTTTGGATATCTAAACATCGTAACCCAGANA 900
 Qy 901 GAAAGGAGACTCACTTATGTGCAATAGCTAGCAACATCGAAGAGATGCCAGATCA 960
 Db 901 GAAAGGAGACTCACTTATGTGCAATAGCTAGCAACATCGAAGAGATGCCAGATCA 960
 Qy 961 GCTGAGTTCGCTACCTTTGTTCTCCAGAGAGATTTGATAGATATTCGGCTATTGGTGG 1020
 Db 961 GCTGAGTTCGCTACCTTTGTTCTCCAGAGAGATTTGATAGATATTCGGCTATTGGTGG 1020
 Qy 1021 TGTCCAAAGCTGAAACAACTCCCACTGGTGGTAAATTTCTTAGAATTTCTATATCAAGAA 1080
 Db 1021 TGTCCAAAGCTGAAACAACTCCCACTGGTGGTAAATTTCTTAGAATTTCTATATCAAGAA 1080
 Qy 1081 AATGATGAATTCGAGGTGGAAATATTTCATGTTACATCCCTCTATGTTGGAAACAAGGAGG 1140

```
Db      ||||| 1081 AATGATGAATCTGAGGTGGAAATTAATTCATGTTACATCCCTATGTTGGAAACAAGGAGG 1140
Qy      ||||| 1141 GCAGATTCATTTCCGTTATCTCTAAACAGGTACAGCAATCCTAAAGTCACTTTTAAAGATG 1200
Db      ||||| 1141 GCAGATTCATTTCCGTTATCTCTAAACAGGTACAGCAATCCTAAAGTCACTTTTAAAGATG 1200
Qy      ||||| 1201 TCAGAAATPAATGATGATGCTGGAAGGAGATCATAGATGTCATAGATGAAGAACTAATTT 1260
Db      ||||| 1201 TCAGAAATPAATGATGATGCTGGAAGGAGATCATAGATGTCATAGATGAAGAACTAATTT 1260
Qy      ||||| 1261 CAACCTTTTGAGATTCATTTTGAAGGAGTTGAAATATATATGSCAGAGCTGATGAGACTCT 1320
Db      ||||| 1261 CAACCTTTTGAGATTCATTTTGAAGGAGTTGAAATATATATGSCAGAGCTGATGAGACTCT 1320
Qy      ||||| 1321 GAGGAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG 1380
Db      ||||| 1321 GAGGAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG 1380
Qy      ||||| 1381 TTGATCTCACCTGAAATTAATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC 1440
Db      ||||| 1381 TTGATCTCACCTGAAATTAATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC 1440
Qy      ||||| 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACAACAGACATC 1500
Db      ||||| 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACAACAGACATC 1500
Qy      ||||| 1501 TGGATAATATCCATGACATCTTTCATGTTTTCCTCCCAAGTCAAGAGGAGAAATTGAG 1560
Db      ||||| 1501 TGGATAATATCCATGACATCTTTCATGTTTTCCTCCCAAGTCAAGAGGAGAAATTGAG 1560
Qy      ||||| 1561 TTTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTAACATCTATT 1620
Db      ||||| 1561 TTTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTAACATCTATT 1620
Qy      ||||| 1621 TTAAGGAAAGCAAAATPAAACGATCCAGTGGTGGCTGCTCTCAAGTGAATTTCAAG 1680
Db      ||||| 1621 TTAAGGAAAGCAAAATPAAACGATCCAGTGGTGGCTGCTCTCAAGTGAATTTCAAG 1680
Qy      ||||| 1681 TGTCTCTACAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTCTGCGCGGAT 1740
Db      ||||| 1681 TGTCTCTACAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTCTGCGCGGAT 1740
Qy      ||||| 1741 GGAATCAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Db      ||||| 1741 GGAATCAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Qy      ||||| 1801 TCCCTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAGG 1860
Db      ||||| 1801 TCCCTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAGG 1860
Qy      ||||| 1861 CTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTCAGACCTGTGACTTTTATA 1920
Db      ||||| 1861 CTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTCAGACCTGTGACTTTTATA 1920
Qy      ||||| 1921 AGTAAGTATAGTAACCAAGAAATCCAACCTGTGTCCTTTTGAAGTATCAAGTCT 1980
Db      ||||| 1921 AGTAAGTATAGTAACCAAGAAATCCAACCTGTGTCCTTTTGAAGTATCAAGTCT 1980
Qy      ||||| 1981 GAAAGTACCCAACTTGCAAAACAAAGAAATTTTGGGCCACCAATTTTGAATTCAGCAGGT 2040
Db      ||||| 1981 GAAAGTACCCAACTTGCAAAACAAAGAAATTTTGGGCCACCAATTTTGAATTCAGCAGGT 2040
Qy      ||||| 2041 CCTCTCTGACTATPACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTACA 2100
Db      ||||| 2041 CCTCTCTGACTATPACTCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTACA 2100
Qy      ||||| 2101 TTGATGGAGTCTCTAAGCCTCATGATCTACAGCTGGAAGAAATATCTCTACTGTG 2160
Db      ||||| 2101 TTGATGGAGTCTCTAAGCCTCATGATCTACAGCTGGAAGAAATATCTCTACTGTG 2160
Qy      ||||| 2161 CTGTTTCATATATGTTGGTCTCTCAGGTGCAGTGGTGAATATCGGTTTAAAGAGTCAAG 2220
```

```
Db      ||||| 2161 CTGTTTCATATATGTTGGTCTCTCAGGTGCAGTGGTGAATTAATCGGTTTAAAGAGTCAAG 2220
Qy      ||||| 2221 TATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGGTTAGTAGACAAACAGG 2280
Db      ||||| 2221 TATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGGTTAGTAGACAAACAGG 2280
Qy      ||||| 2281 GGAATCTGTACCCGAGGGCTTAAATTTGAAGGGCCCTTTAAATATAATAATAAATGSGTCAATA 2340
Db      ||||| 2281 GGAATCTGTACCCGAGGGCTTAAATTTGAAGGGCCCTTTAAATATAATAATAAATGSGTCAATA 2340
Qy      ||||| 2341 GAAATTGACCATCAGTGGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Db      ||||| 2341 GAAATTGACCATCAGTGGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Qy      ||||| 2401 TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGAGATACCTCTCCCTGATGGCATTTA 2460
Db      ||||| 2401 TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGAGATACCTCTCCCTGATGGCATTTA 2460
Qy      ||||| 2461 ATGCAAGGTGAGATATCTTCAAGGTTGCTTATGCTGGGGCCCGAGTCACTCTCTGTGATC 2520
Db      ||||| 2461 ATGCAAGGTGAGATATCTTCAAGGTTGCTTATGCTGGGGCCCGAGTCACTCTCTGTGATC 2520
Qy      ||||| 2521 TTCTATGATACAGGATACACGGAACTTATATGAGGTCACTCTGACACCAAGATGAACAGGGC 2580
Db      ||||| 2521 TTCTATGATACAGGATACACGGAACTTATATATGAGGTCACTCTGACACCAAGATGAACAGGGC 2580
Qy      ||||| 2581 TATTACTTAGGATCTGTGSCCATGCAAGCAGAAAAGTTCCCTCTGTAACCAAAATCGTTTA 2640
Db      ||||| 2581 TATTACTTAGGATCTGTGSCCATGCAAGCAGAAAAGTTCCCTCTGTAACCAAAATCGTTTA 2640
Qy      ||||| 2641 CTGCTCTTACATGTTTCTCTGGATGAGAAATGCCATTTTGGACATACCAAGTATATTACTG 2700
Db      ||||| 2641 CTGCTCTTACATGTTTCTCTGGATGAGAAATGCCATTTTGGACATACCAAGTATATTACTG 2700
Qy      ||||| 2701 AGTTTTTTAGTCAGGGCTGGAAAGCCATATGATTACAGATCTATCTCTCAGGAGACAC 2760
Db      ||||| 2701 AGTTTTTTAGTCAGGGCTGGAAAGCCATATGATTACAGATCTATCTCTCAGGAGACAC 2760
Qy      ||||| 2761 AGCATAAAGATTTCTGAAATCGGGGAAACATTTAGAACTGCACTCTTTTGGCACTACCTTCAA 2820
Db      ||||| 2761 AGCATAAAGATTTCTGAAATCGGGGAAACATTTAGAACTGCACTCTTTTGGCACTACCTTCAA 2820
Qy      ||||| 2821 GAAACCTTGGATACGTTATGCTGCTCTTAAAGATGATTAATTTTGAACCTGTGTAGAAC 2880
Db      ||||| 2821 GAAACCTTGGATACGTTATGCTGCTCTTAAAGATGATTAATTTTGAACCTGTGTAGAAC 2880
Qy      ||||| 2881 TCTCTGTATACACTGGCTTATTTAAACAAATGAGAGGTTTAAATCAACAGAAAACACAGA 2940
Db      ||||| 2881 TCTCTGTATACACTGGCTTATTTAAACAAATGAGAGGTTTAAATCAACAGAAAACACAGA 2940
Qy      ||||| 2941 ATTGATCATCAATTTTGATACCTGCGCATGTAAACATCTACTCTGAAAAATAAATGTGGTG 3000
Db      ||||| 2941 ATTGATCATCAATTTTGATACCTGCGCATGTAAACATCTACTCTGAAAAATAAATGTGGTG 3000
Qy      ||||| 3001 CCAATGAGGGTCTAGGTTTGGTGTAGTAAATCTAAATACCTTAAACCCACATGCTCAAAA 3060
Db      ||||| 3001 CCAATGAGGGTCTAGGTTTGGTGTAGTAAATCTAAATACCTTAAACCCACATGCTCAAAA 3060
Qy      ||||| 3061 TCAATGATACATATTTCTGTGAGAGCCAGCAATACCAATAGAAATTTACTAAAAAATAAATA 3120
Db      ||||| 3061 TCAATGATACATATTTCTGTGAGAGCCAGCAATACCAATAGAAATTTACTAAAAAATAAATA 3120
```

RESULT 3

```
CS196102
LOCUS      3127 bp      DNA      linear      PAT 08-DEC-2005
DEFINITION Sequence 1 from Patent WO2005106021.
ACCESSION CS196102
VERSION    CS196102.1 GI:83410068
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	
-----------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--

3001 CCATCGAGGGTCTACGGTTTGTGTAGTAATCTAATACCTTAACCCCAATGCTCAAAA 3060
3061 TCAATGATACATATTTCTTGAGAGACCAGCAATACCATAGAAATTAATAAAAAA 3120
3061 TCAATGATACATATTTCTTGAGAGACCAGCAATACCATAGAAATTAATAAAAAA 3120
RESULT 4
AX342633 3106 bp DNA linear PAT 12-JAN-2002
LOCUS Sequence 30 from Patent WO0198468.
DEFINITION AX342633
ACCESSION AX342633
VERSION AX342633.1 GI:18152030
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Mammalia; Euarctomylres; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 Yue, H., Elliott, V.S., Gandhi, A.R., Lal, P., Au-Young, J.,
Tribouley, C.M., Deleage, A.M., Baughn, M.R., Nguyen, D.B., Lee, E.A.,
Hafalia, A., Khan, F.A., Wallia, N.K., Iao, M.G., Lu, D.A., Patterson, C.,
Tang, Y.T., Walsh, R.T., Azimzai, Y., Ramkumar, J., Xu, Y. and Reddy, R.
Patent: WO 0198468-A 30 27-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1. 3106
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7160544CB1"
ORIGIN
Query Match 96.2%; Score 3003; DB 2; Length 3106;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3103; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 13 CTCGAGGCGCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGC CGCGCTTTCGCGCT 72
DB 2 CTCGAGGCGCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGC CGCGCTTTCGCGCT 61
QY 73 GGGTTGTCAACGCGCGCGCGCGCGAGGAGGCACTGCAACAGGAGCGGAGTGGAGCGG 132
DB 62 GGGTTGTCAACGCGCGCGCGCGCTGAGGAGGCACTGCAACAGGAGCGGAGTGGAGCGG 121
QY 133 CGCAGCATGAAGCGCGCGCGCGCTGCTAGGAGCGGAGTGGAGCGGAGTGGAGCGG 192
DB 122 CGCAGCATGAAGCGCGCGCGCGCTGCTAGGAGCGGAGTGGAGCGGAGTGGAGCGG 181
QY 193 CGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 252
DB 182 CGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241
QY 253 ATATTTGAACTGCGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 312
DB 242 ATATTTGAACTGCGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 301
QY 313 CTTTTTATGTTGAGCGGTTTCTTCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 372
DB 302 CTTTTTATGTTGAGCGGTTTCTTCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 361
QY 373 AAATATCATGGCTACATGATGCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 432
DB 362 AAATATCATGGCTACATGATGCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 421
QY 433 GATCCAGATGGACCTCATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 492
DB 422 GATCCAGATGGACCTCATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
QY 493 GAAATATACACTGTTTATTTCTGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCG 552
DB 482 GAAATATACACTGTTTATTTCTGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCG 541

1921 AGTAAGTATAGTAACCAAGAGATCCACACTGTGTGCTTTTACAAAGCTATCAAGTCT 1980
1981 GAAGATGACCACTTGGCAAAACAAAGAAATTTTGGGCGACCAATTTGGATTCAGCAGGT 2040
1981 GAAGATGACCACTTGGCAAAACAAAGAAATTTTGGGCGACCAATTTGGATTCAGCAGGT 2040
2041 CCTCTCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTACA 2100
2041 CCTCTCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTACA 2100
2101 TTGTATGGGATGCTCTACAGGCTCATGATCTACAGGCTGGAAGAAATTCCTACTGTG 2160
2101 TTGTATGGGATGCTCTACAGGCTCATGATCTACAGGCTGGAAGAAATTCCTACTGTG 2160
2161 CTGTTCAATATATGTTGCTCTCAGTGAGTGGTGAATATCGGTTTAAAGAGTCAAG 2220
2161 CTGTTCAATATATGTTGCTCTCAGTGAGTGGTGAATATCGGTTTAAAGAGTCAAG 2220
2221 TATTTCCGCTTGAATACCTAGGCTCTAGTGTATGTTGTTAGTATAGAGTCAAG 2280
2221 TATTTCCGCTTGAATACCTAGGCTCTAGTGTATGTTGTTAGTATAGAGTCAAG 2280
2281 GGATCTCTGTCAACGAGGCTTAAATTTGAAGCGCTTTAAATATAAATGGGTCAATA 2340
2281 GGATCTCTGTCAACGAGGCTTAAATTTGAAGCGCTTTAAATATAAATGGGTCAATA 2340
2341 GAAATTTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
2341 GAAATTTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
2401 TTAGATCTGTGGGATCTCAACGCTGCTCTATGAGGAGGAGGAGGAGGAGGAGGAG 2460
2401 TTAGATCTGTGGGATCTCAACGCTGCTCTATGAGGAGGAGGAGGAGGAGGAGGAG 2460
2461 ATGAGAGGTCAGATATCTTTGAGGTCATTTGCTGGGCGGCGGAGTCACTCTGTCGATC 2520
2461 ATGAGAGGTCAGATATCTTTGAGGTCATTTGCTGGGCGGCGGAGTCACTCTGTCGATC 2520
2521 TTCTATGATACAGGATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580
2521 TTCTATGATACAGGATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580
2581 TATTTACTTATGATCTGTGGGATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
2581 TATTTACTTATGATCTGTGGGATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
2641 CTGCTCTTATGATCTGTGGGATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2700
2641 CTGCTCTTATGATCTGTGGGATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2700
2701 AGTTTTTTAGTGGGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2760
2701 AGTTTTTTAGTGGGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2760
2761 AGCATAAAGGTTCTCTGAAATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820
2761 AGCATAAAGGTTCTCTGAAATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820
2821 GAAACCTTTGGATCAGTATGCTGCTCTAAAGTGAATATTTTGAACCTGTGTAGAAC 2880
2821 GAAACCTTTGGATCAGTATGCTGCTCTAAAGTGAATATTTTGAACCTGTGTAGAAC 2880
2881 TCTCTGGTATACACTGGCTTATTTAAACAAATGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940
2881 TCTCTGGTATACACTGGCTTATTTAAACAAATGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940
2941 ATTGATCATCACAATTTTGTATCTGCCATGTAAACATCTACTCTCTGAAATATTAATGTTG 3000
2941 ATTGATCATCACAATTTTGTATCTGCCATGTAAACATCTACTCTCTGAAATATTAATGTTG 3000
3001 CCATGAGGGGCTACGGTTTGTGTAGTATCTTAATCTTAACCTTAACCCCAATGCTCAAAA 3060

Qy 553 CTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCACACCTGGACTATGGAATCTATTCT 612
Db 542 CTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCACACCTGGACTATGGAATCTATTCT 601
Qy 613 CGAAGAAGAAGAACTATTAAAGAGAAAGAAACGCAATTCGAAACAGTCGGAATGCTTTTAC 672
Db 602 CGAAGAAGAAGAACTATTAAAGAGAAAGAAACGCAATTCGAAACAGTCGGAATGCTTTTAC 661
Qy 673 GATTATACCAAGGAGTGGACATCTCTGTTTTCAGCCGCTAGTGGAAATTTATACGTPA 732
Db 662 GATTATACCAAGGAGTGGACATCTCTGTTTTCAGCCGCTAGTGGAAATTTATACGTPA 721
Qy 733 AAAGATGGAGGCGCACAGGATTTACGCCAACACCTTTAAAGGCCCAATCTAGTGGAAACT 792
Db 722 AAAGATGGAGGCGCACAGGATTTACGCCAACACCTTTAAAGGCCCAATCTAGTGGAAACT 781
Qy 793 AGTTGTCGCCACATACGGATCGATCCAAAATTTATGCCCCGCTGATCCAGACTGGATTGCT 852
Db 782 AGTTGTCGCCACATACGGATCGATCCAAAATTTATGCCCCGCTGATCCAGACTGGATTGCT 841
Qy 853 TTTATACATAGCAACGATATTTTGGATATCTAAACATCGTAAACAGAGAAAGAGAGATC 912
Db 842 TTTATACATAGCAACGATATTTTGGATATCTAAACATCGTAAACAGAGAAAGAGAGATC 901
Qy 913 ACTTATGTCACATAGAGTAGCCAAACATGGAAGAAAGATGCCAGATCAGCTGGAGTCGCT 972
Db 902 ACTTATGTCACATAGAGTAGCCAAACATGGAAGAAAGATGCCAGATCAGCTGGAGTCGCT 961
Qy 973 ACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGGCTATTGGTGGTGTCCAAAAGCT 1032
Db 962 ACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGGCTATTGGTGGTGTCCAAAAGCT 1021
Qy 1033 GAAACACCTCCAGTGGTGTGTAATAATCTTAGAAATTTCTATATGAAGAAATATGATGATCT 1092
Db 1022 GAAACACCTCCAGTGGTGTGTAATAATCTTAGAAATTTCTATATGAAGAAATATGATGATCT 1081
Qy 1093 GAGGTGGAATATTTTCATGTTTACATCCCTATGTTGGAAACAGAGAGGCGAGATTCATTC 1152
Db 1082 GAGGTGGAATATTTTCATGTTTACATCCCTATGTTGGAAACAGAGAGGCGAGATTCATTC 1141
Qy 1153 CGTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGTTCAGAAATATG 1212
Db 1142 CGTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGTTCAGAAATATG 1201
Qy 1213 ATTGATGCTGAAGAGGATCATAGATGTCATAGATTAAGGACTAATTCACCTTTTGAG 1272
Db 1202 ATTGATGCTGAAGAGGATCATAGATGTCATAGATTAAGGACTAATTCACCTTTTGAG 1261
Qy 1273 ATTCTATTGAAGGAGTTGAATATATTCAGAGAGTGGATGCACTCTCGAGGGAATAT 1332
Db 1262 ATTCTATTGAAGGAGTTGAATATATTCAGAGAGTGGATGCACTCTCGAGGGAATAT 1321
Qy 1333 GCTTGGTCCATCTTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCT 1392
Db 1322 GCTTGGTCCATCTTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCT 1381
Qy 1393 GAATTATTTATCCAGTAGAAGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTG 1452
Db 1382 GAATTATTTATCCAGTAGAAGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTG 1441
Qy 1453 CTTGATCTGTGAGCGCACTAAATATCTATGAAGAAACAAACAGACATCTGGATAAATATC 1512
Db 1442 CTTGATCTGTGAGCGCACTAAATATCTATGAAGAAACAAACAGACATCTGGATAAATATC 1501
Qy 1513 CATGACATCTTTCAATGTTTTTCCCAAAAGTCACGAAGAGGAAATTTGAGTTTATTTTGGC 1572
Db 1502 CATGACATCTTTCAATGTTTTTCCCAAAAGTCACGAAGAGGAAATTTGAGTTTATTTTGGC 1561
Qy 1573 TCTGAATGCAAAACAGGTTTCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGAC 1632
Db 1562 TCTGAATGCAAAACAGGTTTCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGAC 1621

Qy 1633 AAATATAAACGATCCAGTGGTGGCTGCTGCTCAAGTGATTTCAAGTGTCTCTATCAAA 1692
Db 1622 AAATATAAACGATCCAGTGGTGGCTGCTGCTCAAGTGATTTCAAGTGTCTCTATCAAA 1681
Qy 1693 GAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTCTTGGCCGGCATGATCTAATATC 1752
Db 1682 GAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTCTTGGCCGGCATGATCTAATATC 1741
Qy 1753 CAAAGTTGATGAAGTCCAGAAAGCTGTATATTTTCAAGSCCAACCAAGACTCCCTTTAGAG 1812
Db 1742 CAAAGTTGATGAAGTCCAGAAAGCTGTATATTTTCAAGSCCAACCAAGACTCCCTTTAGAG 1801
Qy 1813 CATCACTCTAGTGTAGTCAAGTTCAGTAAATCCCTGGAGAGGTGACAGGCTGACCGGT 1872
Db 1802 CATCACTCTAGTGTAGTCAAGTTCAGTAAATCCCTGGAGAGGTGACAGGCTGACCGGT 1861
Qy 1873 GGCTACTCACATTTCTTGGCTGCATCAGTCAAGCTGTGCTCCCTTTACAGCTATCAAGTCTGAAAGATGATAGT 1932
Db 1862 GGCTACTCACATTTCTTGGCTGCATCAGTCAAGCTGTGCTCCCTTTACAGCTATCAAGTCTGAAAGATGATAGT 1921
Qy 1933 AACCCAGAAAGATCCACACTGTGCTCCCTTTACAGCTATCAAGTCTGAAAGATGATAGT 1992
Db 1922 AACCCAGAAAGATCCACACTGTGCTCCCTTTACAGCTATCAAGTCTGAAAGATGATAGT 1981
Qy 1993 ACTTGCRAAACAAAGGAATTTTGGGCCCAACATTTTGGATTCAGCAGGTCTCTTCTGAC 2052
Db 1982 ACTTGCRAAACAAAGGAATTTTGGGCCCAACATTTTGGATTCAGCAGGTCTCTTCTGAC 2041
Qy 2053 TATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATG 2112
Db 2042 TATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATG 2101
Qy 2113 CTCTAACAGCTCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTCATATAT 2172
Db 2102 CTCTAACAGCTCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTCATATAT 2161
Qy 2173 GGTGGTCTCAGGTGCAAGTGTGGTGAATTAATCGGTTTAAAGAGAGTCAAGTATTTCCGCTTG 2232
Db 2162 GGTGGTCTCAGGTGCAAGTGTGGTGAATTAATCGGTTTAAAGAGAGTCAAGTATTTCCGCTTG 2221
Qy 2233 AATAACCTTAGCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCTCTGTAC 2292
Db 2222 AATAACCTTAGCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCTCTGTAC 2281
Qy 2293 CGAGGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAATAGAAATTGACGAT 2352
Db 2282 CGAGGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAATAGAAATTGACGAT 2341
Qy 2353 CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTG 2412
Db 2342 CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTG 2401
Qy 2413 GGCATCCAGCGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGGTCA 2472
Db 2402 GGCATCCAGCGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGGTCA 2461
Qy 2473 GATATCTTCAGGGTTCCTATTCCTGGGCCCGAGTCACTCTGTGGATCTTCTATGATACA 2532
Db 2462 GATATCTTCAGGGTTCCTATTCCTGGGCCCGAGTCACTCTGTGGATCTTCTATGATACA 2521
Qy 2533 GGATACCGGAAGGTTATAGGTCACCGCTGACCGAATGAACAGGGCTATTACTAGGA 2592
Db 2522 GGATACCGGAAGGTTATAGGTCACCGCTGACCGAATGAACAGGGCTATTACTAGGA 2581
Qy 2593 TCTGTGCCATCGACAGCAGAAAAGTTCCTCTGAAACCAATCGTTTACTGCTTTACAT 2652
Db 2582 TCTGTGCCATCGACAGCAGAAAAGTTCCTCTGAAACCAATCGTTTACTGCTTTACAT 2641
Qy 2653 GGTTCCTCGATGAGAAATGTCATTTTGGCAATACCAAGTATATTAATGAGTTTTTGTG 2712
Db 2642 GGTTCCTCGATGAGAAATGTCATTTTGGCAATACCAAGTATATTAATGAGTTTTTGTG 2701
Qy 2713 AGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAAGTT 2772

Db 2702 AGGCTGGAAGCCATATGATTTACAGATCTATCTCAGAGAGACACAGCATAGAGTT 2761
Qy 2773 CCTGAATCGGAGAGAACATATGAACTGCACTCTTTTGGCACTACCTTCAAGAAAACTTTGGA 2832
Db 2762 CCTGAATCGGAGAGAACATATGAACTGCACTCTTTTGGCACTACCTTCAAGAAAACTTTGGA 2821
Qy 2833 TCACGTATTGCTGCTCTTAAAGTGATATAATTTGACCTGTGTAGAACTCTCTGTATAC 2892
Db 2822 TCACGTATTGCTGCTCTTAAAGTGATATAATTTGACCTGTGTAGAACTCTCTGTATAC 2881
Qy 2893 ACTGGCTATTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCAC 2952
Db 2882 ACTGGCTATTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCAC 2941
Qy 2953 ATTTTGATACCTGCGCATGTAACTATCTACCTGAAATAATTTGCGTGCATGACGGGT 3012
Db 2942 ATTTTGATACCTGCGCATGTAACTATCTACCTGAAATAATTTGCGTGCATGACGGGT 3001
Qy 3013 CTACGGTTTGGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAATGATAC 3072
Db 3002 CTACGGTTTGGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAATGATAC 3061
Qy 3073 TATTCTGAGAGACCCAGCAATACCAATAGAAATTTACTAAAAAAA 3117
Db 3062 TATTCTGAGAGACCCAGCAATACCAATAGAAATTTACTAAAAAAA 3106

RESULT 5
AX354793
LOCUS AX354793 3143 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 1 from Patent WO0179473.
ACCESSION AX354793
VERSION AX354793.1 GI:18619526
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1
Meyers, R.A. and Williamson, M.
21953, a human prolyl oligopeptidase family member and uses thereof
Patent: WO 0179473-A 1 25-OCT-2001.
Millennium Pharmaceuticals, Inc. (US)

Location/Qualifiers
1..3143
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
229..2877
/note="unnamed protein product"

/codon_start=1
/protein_id="CAD22990.1"
/db_xref="GI:18619527"
/translation="MAAMETEQLGVRIPTADCEINIESQDRPKLPYPPYVERYSWSQ
LKLLADTRKHGYMMAKAPHDPMFKVNDPDGPHSDRIYLYLWAMSGENENTLFSYSEI
PKTINRAAVMLSWKPLDLQATLDYGMTSREELLREKRIGTGVGLASVDYHQSG
TFLQAGSGIYHVDGGPQGTQOPLPNLVSTPCNINRMDPKLCPADPDWIAFLHSN
DWIGNIVTRERRLTYVHNELANMEEDARSAGVATFVLQEEEDFYSGYWCPKAETT
PSGKILRIYVLEENDESVEIHHVTSPLMETRADSPRYPKGTANPKVTPKWSIMI
DAEGRILIDVDEKELQPEILFEGVEVIARAGWTPEGKYAWSILLDRSOTRLQIVLIS
PELFIPEDDWNEROLIESVPDSVTPLIYEETDIWINIHDIHFVFPQSHERRIEF
IFASCKTGFRHLYKIITSLKESYKSSGGLPAPDFKCPKEEIAITSGSEVILGR
HGSNIQVDEVRRLVYFEGKDSPLEHLYVSVYVNPGEVTRLDYDRGYSHSCISQHCD
FTISKYSNQKPHCVSLKLSPPEDPTCKTFEWMATILDSAGLPDPTYPPIFSFES
TTGFTLYGMKLPDLPQKKYPTVFLYGGPQVQLVNNRFKGVKFRILTSLASLYV
VVVIDNRCSGHGLKFECAFVKYKMGQIEIDDOVEGLQYASRYDFDLDRVGHGWSY
GGTSLMALMQRSDFPRVAIAGAPVTLWIFDYGTERYMGPQNEQGYLGSVAMQ
AKKFPSEPNRLLHLGLDENVHFATHSILLISFLVRACKPYDLQIYQBRHSIRVPS
GBHYELHLHYLQENLGSRIALVKVI"

ORIGIN

Query Match 94.8%; Score 2957; DB 2; Length 3143;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 11 GCCTCCGAGGCGCAAGCGCGCTGCTACTGCGCGCGCTCTTCTTAGTGCOCGCTTCGCCGC 70
Db 26 GCCTCCGAGGCGCAAGCGCGCTGCTACTGCGCGCGCTCTTCTTAGTGCOCGCTTCGCCGC 85
Qy 71 CTGGTGTGTCAACCAGCG 130
Db 86 CTGGTGTGTCAACCAGCG 145
Qy 131 GCGCAGCATGAAGCG 190
Db 146 GCGCAGCATGAAGCG 205
Qy 191 GCGCGGGGGAAGGAAATGCAACATGCGCAGCAATGGAACAGAAACAGAAACAGCTGGGTGTG 250
Db 206 GCGCGGGGGAAGGAAATGCAACATGCGCAGCAATGGAACAGAAACAGAAACAGCTGGGTGTG 265
Qy 251 AGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGGCCCTAAATTGG 310
Db 266 AGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGGCCCTAAATTGG 325
Qy 311 AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAAAGCTTAAAGCTTAAAGTAA 370
Db 326 AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAAAGCTTAAAGCTTAAAGTAA 385
Qy 371 GAAATATCATGGCTACATGATGCTAAGGCACCATGATTTTCATGTTTGTGAAGAGA 430
Db 386 GAAATATCATGGCTACATGATGCTAAGGCACCATGATTTTCATGTTTGTGAAGAGA 445
Qy 431 ATGATCCAGATGGACCTCAITTCAGACAGAAATCTATTACCTTGGCCATGTCTGGTGAA 490
Db 446 ATGATCCAGATGGACCTCAITTCAGACAGAAATCTATTACCTTGGCCATGTCTGGTGAA 505
Qy 491 GAGAAAACTACTGTTTATTTCTGAAATTTCCAAAACATCAATAGAGAGAGCTCTTAA 550
Db 506 GAGAAAACTACTGTTTATTTCTGAAATTTCCAAAACATCAATAGAGAGAGCTCTTAA 565
Qy 551 TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGGACTATGGAATGTATT 610
Db 566 TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGGACTATGGAATGTATT 625
Qy 611 CTCGAGAGAGAACTATTAAGAGAGAGAAAAACCATTTGGAAACAGTCCGAAATTTGCTTCT 670
Db 626 CTCGAGAGAGAACTATTAAGAGAGAGAAAAACCATTTGGAAACAGTCCGAAATTTGCTTCT 685
Qy 671 ACATTTATCAACAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCACG 730
Db 686 ACATTTATCAACAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCACG 745
Qy 731 TAAAGATCGAGGCGCAACAGGATTTAGCAACAACTTTTAAAGGCCCAATCTAGTGGAAA 790
Db 746 TAAAGATCGAGGCGCAACAGGATTTAGCAACAACTTTTAAAGGCCCAATCTAGTGGAAA 805
Qy 791 CTAGTTTGTCCCAACATACGATGAGTTCGAAATTTATGCCCCGTGATCCAGACTGGATTG 850
Db 806 CTAGTTTGTCCCAACATACGATGAGTTCGAAATTTATGCCCCGTGATCCAGACTGGATTG 865
Qy 851 CTTTTATATACATAGCAACGATTTTGGATATCTTACATCGTAAACAGAGAGAGAGAGAGAC 910
Db 866 CTTTTATATACATAGCAACGATTTTGGATATCTTAAACATCGTAAACAGAGAGAGAGAGAGAC 925
Qy 911 TCATTTATGTGCAATAGCTAGTACCAACATCGAAGAGAGATGCCAGATCAGCTGGAGTCG 970
Db 926 TCATTTATGTGCAATAGCTAGTACCAACATCGAAGAGAGATGCCAGATCAGCTGGAGTCG 985
Qy 971 CTACCTTTTGTCTTCAAGAGAAATTTGATAGATATTTGGCTATTGGGTGTCTCAAAAG 1030
Db 986 CTACCTTTTGTCTTCAAGAGAAATTTGATAGATATTTGGCTATTGGGTGTCTCAAAAG 1045
Qy 1031 CTGAAACAACTCCAGCTGGTGTGTAATAATTTCTTAGAATTTCTATATGAAGAAAAATGATGAAT 1090

Db 1046 |||||CTGAACAATCCCAAGTGGTAAATCTTAGAATCTATATGAAGAAATGATGAAT 1105
Qy 1091 CTGAGTGGAAATATTATCATGTTACATCCCTATGTTGGAAAACAGAGAGGCGAGATTCAAT 1150
Db 1106 CTGAGGTGGAAATATTATCATGTTACATCCCTATGTTGGAAAACAGAGAGGCGAGATTCAAT 1165
Qy 1151 TCCGTTATCTTAAACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATAA 1210
Db 1166 TCCGTTATCTTAAACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATAA 1225
Qy 1211 TGAATGATGCTGAAAGAGGATCATAGATGTCATAGATAAGAACTAAATTCACCTTTTG 1270
Db 1226 TGAATGATGCTGAAAGAGGATCATAGATGTCATAGATAAGAACTAAATTCACCTTTTG 1285
Qy 1271 AGATTCTATTTGAAAGAGTTGAATATATTTGCAGAGCTGGATGGACTCTCTGAGGGAAAT 1330
Db 1286 AGATTCTATTTGAAAGAGTTGAATATATTTGCAGAGCTGGATGGACTCTCTGAGGGAAAT 1345
Qy 1331 ATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAC 1390
Db 1346 ATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCCTGCAGATAGTGTGATCTCAC 1405
Qy 1391 CTGAATATTATTTCCAGTAGAAGATGATGTTATGGAAAGGCGAGAGACTCATTTGAGTCAG 1450
Db 1406 CTGAATATTATTTCCAGTAGAAGATGATGTTATGGAAAGGCGAGAGACTCATTTGAGTCAG 1465
Qy 1451 TGCCTGATTCTGTGAGCGCACTTAATTTATCTATGAAGAAACAACAGACATCTGGATANATA 1510
Db 1466 TGCCTGATTCTGTGAGCGCACTTAATTTATCTATGAAGAAACAACAGACATCTGGATANATA 1525
Qy 1511 TCCATGACATCTTTTCATGTTTTTCCCAAGTGCAGAGGAGAAATGAGTTATTTTGG 1570
Db 1526 TCCATGACATCTTTTCATGTTTTTCCCAAGTGCAGAGGAGAAATGAGTTATTTTGG 1585
Qy 1571 CCTCTGAATGCAAAACAGTTTCGGTCAATTTATACAAAATATACATCTATTTTAAAGGAAA 1630
Db 1586 CCTCTGAATGCAAAACAGTTTCGGTCAATTTATACAAAATATACATCTATTTTAAAGGAAA 1645
Qy 1631 GCAATATTAACAGTACAGTGGTGGCTGCTGCTCAAGTGATTTCAAGTGCTCTATCA 1690
Db 1646 GCAATATTAACAGTACAGTGGTGGCTGCTGCTCAAGTGATTTCAAGTGCTCTATCA 1705
Qy 1691 AAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATA 1750
Db 1706 AAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATA 1765
Qy 1751 TCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAG 1810
Db 1766 TCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAG 1825
Qy 1811 AGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACC 1870
Db 1826 AGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACC 1885
Qy 1871 GTGGCTACTCACATTTCTGCTGCAATCAGTCAGCACTGTGACTTCTTTTATAAGTAAATATA 1930
Db 1886 GTGGCTACTCACATTTCTGCTGCAATCAGTCAGCACTGTGACTTCTTTTATAAGTAAATATA 1945
Qy 1931 GTAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCTGAAATGACC 1990
Db 1946 GTAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCTGAAATGACC 2005
Qy 1991 CAACTTGCAAAACAAAGAAATTTTGGGCCACCAATTTGGATTCAGCAGGTCTCTCTCTG 2050
Db 2006 CAACTTGCAAAACAAAGAAATTTTGGGCCACCAATTTGGATTCAGCAGGTCTCTCTCTG 2065
Qy 2051 ACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGATTTACATTTGATGGGA 2110
Db 2066 ACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGATTTACATTTGATGGGA 2125
Qy 2111 TGCTCTACAGCTCATGATCTACAGCTCGGAAGAAATATCTCTATGTGCTGTTCAAT 2170
|||||

Db 2126 TGCTCTACAAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTCATAT 2185
Qy 2171 ATGGTGGTCTCTCAGGTGCAAGTTGGTGAATAATTCGGTTTAAAGAGAGTCAAGTATTTCCGCT 2230
Db 2186 ATGGTGGTCTCTCAGGTGCAAGTTGGTGAATAATTCGGTTTAAAGAGAGTCAAGTATTTCCGCT 2245
Qy 2231 TGAATACCCCTAGCCTCTCTAGGTTATGTGGTTGTAGTATAGACAACAGAGGGATCTCTGTC 2290
Db 2246 TGAATACCCCTAGCCTCTCTAGGTTATGTGGTTGTAGTATAGACAACAGAGGGATCTCTGTC 2305
Qy 2291 ACCGAGGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATCGGTCAAAATAGAAATTCAGC 2350
Db 2306 ACCGAGGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATCGGTCAAAATAGAAATTCAGC 2365
Qy 2351 ATCAGGTGCAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTG 2410
Db 2366 ATCAGGTGCAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTG 2425
Qy 2411 TGGGCATCCACGGCTGGTCTTATGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATA 2470
Db 2426 TGGGCATCCACGGCTGGTCTTATGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATA 2485
Qy 2471 CAGATATCTTCAAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATA 2530
Db 2486 CAGATATCTTCAAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATA 2545
Qy 2531 CAGGATACACGGAAGCTTATATGCTGCTGACCCCTGACCAAGATGAAACAGGGCTATTTACTTAG 2590
Db 2546 CAGGATACACGGAAGCTTATATGCTGCTGACCCCTGACCAAGATGAAACAGGGCTATTTACTTAG 2605
Qy 2591 GATCTGTGGCCATGCAAGAGAAAGTTTCCCTCTGAACCAAAATCGTTTACTCTCTTAC 2650
Db 2606 GATCTGTGGCCATGCAAGAGAAAGTTTCCCTCTGAAACCAAAATCGTTTACTCTCTTAC 2665
Qy 2651 ATGGTTTCTGTGATGAGAAATGTCATTTTGCACATACAGTATATTTACTCAGTCTTTTCTAG 2710
Db 2666 ATGGTTTCTGTGATGAGAAATGTCATTTTGCACATACAGTATATTTACTCAGTCTTTTCTAG 2725
Qy 2711 TGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGAG 2770
Db 2726 TGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGAG 2785
Qy 2771 TTCTCTGAATTCGGGGAACATTTATGAACTGTCATCTTTTGCACACTACCTTCAAGAAAACTTG 2830
Db 2786 TTCTCTGAATTCGGGGAACATTTATGAACTGTCATCTTTTGCACACTACCTTCAAGAAAACTTG 2845
Qy 2831 GATCACGTATTGCTGCTCTTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGSTAT 2890
Db 2846 GATCACGTATTGCTGCTCTTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGSTAT 2905
Qy 2891 ACACCTGGCTATTAAACCAAAATGAGAGGTTTAAATCAACAGAAACACAGAAATTTGATCATC 2950
Db 2906 ACACCTGGCTATTAAACCAAAATGAGAGGTTTAAATCAACAGAAACACAGAAATTTGATCATC 2965
Qy 2951 ACATTTTGTATACCTGCGCATGTAACATCTACTCTGAAATAAATGTTGGTCCCATGAGGG 3010
Db 2966 ACATTTTGTATACCTGCGCATGTAACATCTACTCTGAAATAAATGTTGGTCCCATGAGGG 3025
Qy 3011 GTCTACGGTTTGTGTAGTAACTTAATACCTTAACCCCATGCTCAAAATCAAAATGATA 3070
Db 3026 GTCTACGGTTTGTGTAGTAACTTAATACCTTAACCCCATGCTCAAAATCAAAATGATA 3085
Qy 3071 CATATTTCTGAGAGACCCAGCAATACATTAAGAAATTTACTTAAAAAATAAAAAA 3120
Db 3086 CATATTTCTGAGAGACCCAGCAATACATTAAGAAATTTACTTAAAAAATAAAAAA 3135
|||||

RESULT 6
ARG31283
LOCUS
DEFINITION
ACCESSION
VERSION

ARG31283
Sequence 12 from patent US 6944180.
ARG31283
VERSION
4829 bp
DNA
linear
PAT 14-FEB-2005
GI:59770927

KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 4829)
AUTHORS	Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE	Serine protease genes related to DPP1V
JOURNAL	Patent: US 6841180-A 12 18-JAN-2005;
	Ferring BV,;
FEATURES	Location/Qualifiers
source	1..4829
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
	Query Match 92.6%; Score 2889; DB 2; Length 4829;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
Qy	1 AAGTGTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTTTGTAGTCCG 60
Db	1 AAGTGTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTTTGTAGTCCG 60
Qy	61 CGTTCCCGCTGGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTGCAACCAAGACCG 120
Db	61 CGTTCCCGCTGGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTGCAACCAAGACCG 120
Qy	121 GAGTGAAGGCGGCGAGCATGAAGCGGCGAGGCCGCTCCATAGCGCACGTGGGACGG 180
Db	121 GAGTGAAGGCGGCGAGCATGAAGCGGCGAGGCCGCTCCATAGCGCACGTGGGACGG 180
Qy	181 TCCGGCGGCGCGCGGCGAGGAAATGCAACATGCGCAGCAGCAATGGAAACAGAACAG 240
Db	181 TCCGGCGGCGCGCGGCGAGGAAATGCAACATGCGCAGCAGCAATGGAAACAGAACAG 240
Qy	241 CTGGGTGTTGAGATATTTGAACTCGGGACTGTGAGGAGATATTTGAATCAGAGATCGG 300
Db	241 CTGGGTGTTGAGATATTTGAACTCGGGACTGTGAGGAGATATTTGAATCAGAGATCGG 300
Qy	301 CCTAAATGGAGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTTAAAGAGCTGCTT 360
Db	301 CCTAAATGGAGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTTAAAGAGCTGCTT 360
Qy	361 GCCGATACAGAAATATCATGGCTACATGATGGCTTAAGGCAACACATGATTTTATGTTT 420
Db	361 GCCGATACAGAAATATCATGGCTACATGATGGCTTAAGGCAACACATGATTTTATGTTT 420
Qy	421 GTGAAGAGGAATGATTCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCT 480
Db	421 GTGAAGAGGAATGATTCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCT 480
Qy	481 GGTGAGACAGGAAATACACTGTTTATTTCTGAAATCCCAAACTATCAATAGAGCA 540
Db	481 GGTGAGACAGGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Qy	541 GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACATCTGACTAT 600
Db	541 GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACATCTGACTAT 600
Qy	601 GGAATGTATTCTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
Db	601 GGAATGTATTCTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
Qy	661 ATTGCTTCTTACGATTATCACCAAGGAAGTGGAAATTTCTGTTTCAAGCCGCTAGTGA 720
Db	661 ATTGCTTCTTACGATTATCACCAAGGAAGTGGAAATTTCTGTTTCAAGCCGCTAGTGA 720
Qy	721 ATTTATCAGCTAAAGATGAGGGCCACAGGATTTACGCAACAACTTTAAAGGCCAAT 780
Db	721 ATTTATCAGCTAAAGATGAGGGCCACAGGATTTTACGCAACAACTTTTAAAGGCCAAT 780
Qy	781 CTAGTGGAAACTAGTTGTGCCCAACATACGGATGGATCCAAATTTATGCCCGCTGATCCA 840

Db	781 CTAGTGGAAACTAGTTGTGCCCAACATACGGATGGATCCAAATTTATGCCCTGCTGATCCA 840
Qy	841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAAATCGTAAACAGAGAA 900
Db	841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAAATCGTAAACAGAGAA 900
Qy	901 GAAAGGAGACTCAGTTATGTCACAAATGAGCTAGCCAAATGGAAGAAAGTGCAGATCA 960
Db	901 GAAAGGAGACTCAGTTATGTCACAAATGAGCTAGCCAAATGGAAGAAAGTGCAGATCA 960
Qy	961 GCTGGAGTGGCTACCTTTGTTCTCAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020
Db	961 GCTGGAGTGGCTACCTTTGTTCTCAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020
Qy	1021 TGTCCAAAGAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Db	1021 TGTCCAAAGAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Qy	1081 AATGATGAATCTGAGGTGGAATTTATCATGTTACATCCCTATGTTGGAAACAGAGAGG 1140
Db	1081 AATGATGAATCTGAGGTGGAATTTATCATGTTACATCCCTATGTTGGAAACAGAGAGG 1140
Qy	1141 GCAGATTCAATTCGTTATCCTTAAACAGGTACAGCAATCCTAAAGTCACCTTTTAAAGATG 1200
Db	1141 GCAGATTCAATTCGTTATCCTTAAACAGGTACAGCAATCCTTAAAGTCACCTTTTAAAGATG 1200
Qy	1201 TCAGAAATATGATGATGCTGTAAGGAGGATCATAGATGTCATAGATAAGAACTAATTT 1260
Db	1201 TCAGAAATATGATGATGCTGTAAGGAGGATCATAGATGTCATAGATAAGAACTAATTT 1260
Qy	1261 CAACCTTTTTCAGATTTCTTATTTGAAGGAGTTGAATATTTGCCAGAGCTGGATGGACTCCT 1320
Db	1261 CAACCTTTTTCAGATTTCTTATTTGAAGGAGTTGAATATTTGCCAGAGCTGGATGGACTCCT 1320
Qy	1321 GAGGGAATATATGCTTTGGTCCATCCTACTAGATCGCTCCACAGCTCGCTACAGATAGTG 1380
Db	1321 GAGGGAATATATGCTTTGGTCCATCCTACTAGATCGCTCCACAGCTCGCTACAGATAGTG 1380
Qy	1381 TTGATCTCAGCTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC 1440
Db	1381 TTGATCTCAGCTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC 1440
Qy	1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCACTAATTTATCTATGAAAGAAACACAGACATC 1500
Db	1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCACTAATTTATCTATGAAAGAAACACAGACATC 1500
Qy	1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTTCCCAAGTCAAGAGGAAATTTGAG 1560
Db	1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTTCCCAAGTCAAGAGGAAATTTGAG 1560
Qy	1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTACATCTATT 1620
Db	1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTACATCTATT 1620
Qy	1621 TTTAAAGGAAGCAAAATATAAACGATCCAGTGGGCTGCTCTCCAGTGAATTTTCAAG 1680
Db	1621 TTTAAAGGAAGCAAAATATAAACGATCCAGTGGGCTGCTCTCTCCAGTGAATTTTCAAG 1680
Qy	1681 TGTCTCTATCAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Db	1681 TGTCTCTATCAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Qy	1741 GGAATCTAATATCCAAAGTGAAGTCAAGAGGCTGGTATATTTTGAAGGACCAACAGAC 1800
Db	1741 GGAATCTAATATCCAAAGTGAAGTCAAGAGGCTGGTATATTTTGAAGGACCAACAGAC 1800
Qy	1801 TCCCTTTTAGAGCATCAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 1860
Db	1801 TCCCTTTTAGAGCATCAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 1860
Qy	1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGACCTGACCTGACCTGACCTGACCTGACCT 1920

Db	1861	CTGACTGACCGTGGCTACTCACAATCTTGTCTGTCATCAGTCAGCACTGTGACTTCTTTATA	1920
Qy	1921	AGTAAGTATAGTAACACAGAGAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCT	1980
Db	1921	AGTAAGTATAGTAACACAGAGAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCT	1980
Qy	1981	GAAGATGACCCCAACTTGCAGAACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGAGT	2040
Db	1981	GAAGATGACCCCAACTTGCAGAACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGAGT	2040
Qy	2041	CCTCTTCTCGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGAATTTACA	2100
Db	2041	CCTCTTCTCGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGAATTTACA	2100
Qy	2101	TTGTATGGATGCTCTACAGCCCTCATGATCTACAGCTTGGAAAGAAATATCTACTGTG	2160
Db	2101	TTGTATGGATGCTCTACAGCCCTCATGATCTACAGCTTGGAAAGAAATATCTACTGTG	2160
Qy	2161	CTGTTTCATATATGTGG--TCTCAGGTGAGTTGGTGAATATCGTTTAAAGGAGTCA	2218
Db	2161	CTGTTTCATATATGTGGTCTCTCAGGTGAGTTGGTGAATATCGTTTAAAGGAGTCA	2220
Qy	2219	AGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACA	2278
Db	2221	AGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACA	2280
Qy	2279	GGGATCTCTGTACCGAGGGCTTAAATTTGAAGGCCCTTTAAATATATAAATGGGTCAAA	2338
Db	2281	GGGATCTCTGTACCGAGGGCTTAAATTTGAAGGCCCTTTAAATATATAAATGGGTCAAA	2340
Qy	2339	TAGAAATGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATCATTTCAATG	2398
Db	2341	TAGAAATGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATCATTTCAATG	2400
Qy	2399	ACTTAGATCTGTGGGCATCCAGGCTGTCTCTATGGAGGATACCTCTCCCTGATGCAT	2458
Db	2401	ACTTAGATCTGTGGGCATCCAGGCTGTCTCTATGGAGGATACCTCTCCCTGATGCAT	2460
Qy	2459	TAATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGCCCCCACTCTGTGGA	2518
Db	2461	TAATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGCCCCCACTCTGTGGA	2520
Qy	2519	TCTTCTATGATACAGGATACACGAAAGTATATGGGTCAACCTTGCAGCAGATGAAACAG	2578
Db	2521	TCTTCTATGATACAGGATACACGAAAGTATATGGGTCAACCTTGCAGCAGATGAAACAG	2580
Qy	2579	GCTATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCGCTCTGAAACCAATCGTT	2638
Db	2581	GCTATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCGCTCTGAAACCAATCGTT	2640
Qy	2639	TACTGCTCTTACATGTTTCTCTGGATGGAATGTCCATTTTGGACATACCATGATATTAC	2698
Db	2641	TACTGCTCTTACATGTTTCTCTGGATGGAATGTCCATTTTGGACATACCATGATATTAC	2700
Qy	2699	TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAGAC	2758
Db	2701	TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAGAC	2760
Qy	2759	ACAGCATAGAGTTCTGATCGGAGAACATTTATGAATCTGATCTTTTGGCACTACTCTTC	2818
Db	2761	ACAGCATAGAGTTCTGATCGGAGAACATTTATGAATCTGATCTTTTGGCACTACTCTTC	2820
Qy	2819	AGAAACCTTGGATCAGTATGTCTCTTAAAGTGATATAATTTTGAACCTGTGTAGA	2878
Db	2821	AGAAACCTTGGATCAGTATGTCTCTTAAAGTGATATAATTTTGAACCTGTGTAGA	2880
Qy	2879	ACTCTCTGTATACCTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACA	2938
Db	2881	ACTCTCTGTATACCTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACA	2940
Qy	2939	GAATGATCATCACAATTTTGATCCTGCCATGTAAACATCTACTCTCTGAAATAAATGTGG	2998
Db	2941	GAATGATCATCACAATTTTGATCCTGCCATGTAAACATCTACTCTCTGAAATAAATGTGG	3000
Qy	2999	TGCCATGACAGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAACCCACATGCTCAA	3058
Db	3001	TGCCATGACAGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAACCCACATGCTCAA	3060
Qy	3059	AATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCATAAGAAATTAATAAAAAA	3118
Db	3061	AATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCATAAGAAATTAATAAAAAA	3120
Qy	3119	AA 3120	
Db	3121	AA 3122	
RESULT 7			
AX608735			
LOCUS			
DEFINITION			
AX608735			
ACCESSION			
AX608735.1			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eumetazoa; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
1			
Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.			
Novel serine protease genes related to appiv			
Patent: WO 0231134-A 12 18-APR-2002;			
Ferring BV (NL)			
FEATURES			
Location/Qualifiers			
source			
1..4829			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match			
Best Local Similarity 92.6%; Score 2889; DB 2; Length 4829;			
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;			
Qy	1	AAGTGCTTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCCGCTGCTTCTTAGTGCGG	60
Db	1	AAGTGCTTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCCGCTGCTTCTTAGTGCGG	60
Qy	61	CGTTGCGCGCTGGGTTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120
Db	61	CGTTGCGCGCTGGGTTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120
Qy	121	GAGTGGAGGCG	180
Db	121	GAGTGGAGGCG	180
Qy	181	TCCGGGCG	240
Db	181	TCCGGGCG	240
Qy	241	CTGGGTGTTGAGATATTTGAACCTCGGACCTGTGGAGGAGATATTTGAATCAGAGATCGG	300
Db	241	CTGGGTGTTGAGATATTTGAACCTCGGACCTGTGGAGGAGATATTTGAATCAGAGATCGG	300
Qy	301	CCTAATTCGAGCGCTTTTATGTTGAGCGGTATTTCTTGGAGTCAAGCTTAAAAAGCTGCTT	360
Db	301	CCTAATTCGAGCGCTTTTATGTTGAGCGGTATTTCTTGGAGTCAAGCTTAAAAAGCTGCTT	360
Qy	361	GGCGATACAGAAATATCATGGCTTACATGATGCTTAAAGGACCAACATGATTTTCTGTTT	420
Db	361	GGCGATACAGAAATATCATGGCTTACATGATGCTTAAAGGACCAACATGATTTTCTGTTT	420
Qy	421	GTGAGAGGAGATGATCCAGATGGACCTCATTTACAGAGAAATCTATTACCTTGGCATGTCT	480
Db	421	GTGAGAGGAGATGATCCAGATGGACCTCATTTACAGAGAAATCTATTACCTTGGCATGTCT	480

Qy	481	GGTGAACAGAGAAATACACTGTTTATCTGAAATCCCAAACTATCAATAGAGCA	540	1561	TTTATTTTTGCTCTCGAATGATAACAGGTTTCGGTCAATTTATATACAAAATTTACATCTATT	1620
Db	481	GGTGAACAGAGAAATACACTGTTTATCTGAAATCCCAAACTATCAATAGAGCA	540	1561	TTTATTTTTGCTCTCGAATGATAACAGGTTTCGGTCAATTTATATACAAAATTTACATCTATT	1620
Qy	541	GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGAGTCTTTTTCAGGCAACACTGGACTAT	600	1621	TTAAAGGAAAGCAATATATAACGATCCAGTGGTGGGCTGCTCTCAAGTGATTTTCAAG	1680
Db	541	GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGAGTCTTTTTCAGGCAACACTGGACTAT	600	1621	TTAAAGGAAAGCAATATATAACGATCCAGTGGTGGGCTGCTCTCAAGTGATTTTCAAG	1680
Qy	601	GGATGATTTCTCGAGAGAGAACTATTAAGAGAAAGAAACGATTTGGAACAGTCGGA	660	1681	TGTCCTTATCAAGAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTTGGCCGCGCAT	1740
Db	601	GGATGATTTCTCGAGAGAGAACTATTAAGAGAAAGAAACGATTTGGAACAGTCGGA	660	1681	TGTCCTTATCAAGAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTTGGCCGCGCAT	1740
Qy	661	ATTGCTTTCTTACGATTTATCACAAGGAAGTGAACATTTCTGTGTTTCAAGCCGGTAGTGA	720	1741	GGATCTAATATCCAACTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGACCAAGAC	1800
Db	661	ATTGCTTTCTTACGATTTATCACAAGGAAGTGAACATTTCTGTGTTTCAAGCCGGTAGTGA	720	1741	GGATCTAATATCCAACTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGACCAAGAC	1800
Qy	721	ATTTATCAGTAAAGATGGAGGCGCACAGGATTTAGCGAAACACCTTTAAGGCCCAAT	780	1801	TCCCTTTTAGAGCATCACCTGTACGTAAGTACGTAAGTACGTAAGTACGTAAGTACGTAAGT	1860
Db	721	ATTTATCAGTAAAGATGGAGGCGCACAGGATTTAGCGAAACACCTTTAAGGCCCAAT	780	1801	TCCCTTTTAGAGCATCACCTGTACGTAAGTACGTAAGTACGTAAGTACGTAAGTACGTAAGT	1860
Qy	781	CTAGTGAACACTAGTTGTCCTCAACATACGATGGAATCCAAATTTATGCCCCGCTGATCCA	840	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA	1920
Db	781	CTAGTGAACACTAGTTGTCCTCAACATACGATGGAATCCAAATTTATGCCCCGCTGATCCA	840	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA	1920
Qy	841	GACTGGATTCCTTTTATACATAGCAACAGATATTTGGATATCTAACATCGTAACAGAGAA	900	1921	AGTAAGTATAGTAAACAGAGAAATCCACATCTGTGTGCCCTTTTCAAGCTATCAAGTCTCT	1980
Db	841	GACTGGATTCCTTTTATACATAGCAACAGATATTTGGATATCTAACATCGTAACAGAGAA	900	1921	AGTAAGTATAGTAAACAGAGAAATCCACATCTGTGTGCCCTTTTCAAGCTATCAAGTCTCT	1980
Qy	901	GAAAGGAGACTCTTATATGTCACATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA	960	1981	GAAAGTGAACCACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGT	2040
Db	901	GAAAGGAGACTCTTATATGTCACATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA	960	1981	GAAAGTGAACCACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGT	2040
Qy	961	GCTGGAGTCCGTACTTGTCTTCCAGAGAAATTTGATAGATATCTTGGCTATTTGGTGG	1020	2041	CCTCTTCTGATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCGATTTTACA	2100
Db	961	GCTGGAGTCCGTACTTGTCTTCCAGAGAAATTTGATAGATATCTTGGCTATTTGGTGG	1020	2041	CCTCTTCTGATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCGATTTTACA	2100
Qy	1021	TGTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATGAGAA	1080	2101	TTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTCATGATCTACAGCTGGAAGAAATATCTACTGTG	2160
Db	1021	TGTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATGAGAA	1080	2101	TTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTCATGATCTACAGCTGGAAGAAATATCTACTGTG	2160
Qy	1081	AATGATGAATCTGAGTGGAAATTTATCATGTTTACATCCCTTATGTTGGAACAGGAGG	1140	2161	CTGTTTCATATATGTTGG--TCCTCAGTGCAGTGGTGGTGAATATCCGTTTAAAGGAGTCA	2218
Db	1081	AATGATGAATCTGAGTGGAAATTTATCATGTTTACATCCCTTATGTTGGAACAGGAGG	1140	2161	CTGTTTCATATATGTTGG--TCCTCAGTGCAGTGGTGGTGAATATCCGTTTAAAGGAGTCA	2218
Qy	1141	GCAGATTCATTCGGTATCTTAAACAGGTCACAGCAATCTTAAAGTCACTTTTAAAGATG	1200	2219	AGTATTTCCGCTTGAATACCTAGCTCTCTAGTGTATGTTGTTGTAGTATAGACAA	2278
Db	1141	GCAGATTCATTCGGTATCTTAAACAGGTCACAGCAATCTTAAAGTCACTTTTAAAGATG	1200	2219	AGTATTTCCGCTTGAATACCTAGCTCTCTAGTGTATGTTGTTGTAGTATAGACAA	2278
Qy	1201	TCAGAAATTAATGATGCTGGAAGGAGATCATAGATGTCATAGATAAGGAACTAAT	1260	2279	GGGGATCTCTGTACCGAGGCTTAAATTTGAAGGCGCTTAAATATATAAATGGGTCAA	2338
Db	1201	TCAGAAATTAATGATGCTGGAAGGAGATCATAGATGTCATAGATAAGGAACTAAT	1260	2279	GGGGATCTCTGTACCGAGGCTTAAATTTGAAGGCGCTTAAATATATAAATGGGTCAA	2338
Qy	1261	CAACCTTTTGAATTTATTTGAAGGAGTGAATATATGTCAGAGCTGGATGGACTCCT	1320	2339	TAGAAATTTGAAGATCAGTGGAGGACTCCAAATCTAGCTTCCGATATGATTTTCAATTTG	2398
Db	1261	CAACCTTTTGAATTTATTTGAAGGAGTGAATATATGTCAGAGCTGGATGGACTCCT	1320	2339	TAGAAATTTGAAGATCAGTGGAGGACTCCAAATCTAGCTTCCGATATGATTTTCAATTTG	2398
Qy	1321	GAGGAAATATGCTTGTGCTTACATAGATGCTGAGGAGGATCATAGATGTCATAGATAAGGAACTAAT	1380	2401	ACTTAGATCGTGGGCTCCAGGCTGCTCTATGAGGATACCTCTCCCTCATGGCAT	2458
Db	1321	GAGGAAATATGCTTGTGCTTACATAGATGCTGAGGAGGATCATAGATGTCATAGATAAGGAACTAAT	1380	2401	ACTTAGATCGTGGGCTCCAGGCTGCTCTATGAGGATACCTCTCCCTCATGGCAT	2458
Qy	1381	TTGATCTACCTGAAATTTATTTTCCAGTGAAGATGATGTTATGGAAGGCGAGACTC	1440	2459	TAATGAGAGGTCAGATATCTTCAAGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGA	2518
Db	1381	TTGATCTACCTGAAATTTATTTTCCAGTGAAGATGATGTTATGGAAGGCGAGACTC	1440	2459	TAATGAGAGGTCAGATATCTTCAAGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGA	2518
Qy	1441	ATTGAGTCAGTCCGCTGATGAGGCACTAATTTATCTATGAGAAACAGACATC	1500	2519	TCCTTCTATGATACAGGATACAGGAACTTATATGGGTCACTCCCTGACCAAGTCAACAGG	2578
Db	1441	ATTGAGTCAGTCCGCTGATGAGGCACTAATTTATCTATGAGAAACAGACATC	1500	2519	TCCTTCTATGATACAGGATACAGGAACTTATATGGGTCACTCCCTGACCAAGTCAACAGG	2578
Qy	1501	TGATTAATATCCATGACATCTTCTATGTTTCCCAAGTCAAGAGGAGAAATTTGAG	1560	2579	GCTATTTACTTAGGATCTGTGGCATGCAAGCAGAAAGTTCCTCTCTGAAACCAATCGTT	2638
Db	1501	TGATTAATATCCATGACATCTTCTATGTTTCCCAAGTCAAGAGGAGAAATTTGAG	1560	2579	GCTATTTACTTAGGATCTGTGGCATGCAAGCAGAAAGTTCCTCTCTGAAACCAATCGTT	2638
Qy	1501	TGATTAATATCCATGACATCTTCTATGTTTCCCAAGTCAAGAGGAGAAATTTGAG	1560	2639	TACTGCTCTTACATGTTTCTCTGGATGAGATGTCATTTTGGCATACATACCATGATATATAC	2698
Db	1501	TGATTAATATCCATGACATCTTCTATGTTTCCCAAGTCAAGAGGAGAAATTTGAG	1560	2639	TACTGCTCTTACATGTTTCTCTGGATGAGATGTCATTTTGGCATACATACCATGATATATAC	2698

```
Db 2641 TACTGCTTACATGTTCTCTGGATGAGAAATGCCATTTTGCACATACCAGTATATTAC 2700
Qy TGAATTTTCTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCCTCAGGAGAGAC 2758
Db 2701 TGAATTTTCTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCCTCAGGAGAGAC 2760
Qy ACAGCATAGAGTTCTGATTCGGGAGAACATTAAGAAGTGCATCTTTTGCACATCCTTTC 2818
Db 2761 ACAGCATAGAGTTCTGATTCGGGAGAACATTAAGAAGTGCATCTTTTGCACATCCTTTC 2820
Qy AGAAATCCTTGGATCAGTATGCTGCTCTAAAGTGATATAATTTGACCTGTGTAGA 2878
Db 2821 AGAAATCCTTGGATCAGTATGCTGCTCTAAAGTGATATAATTTGACCTGTGTAGA 2880
Qy ACTCTCTGTTATACACTGGCTATTATAACCAATGAGGAGGTTTAAATCAACAGAAACACA 2938
Db 2881 ACTCTCTGTTATACACTGGCTATTATAACCAATGAGGAGGTTTAAATCAACAGAAACACA 2940
Qy GAATTCATCATCACATTTTGTATACCTGCCATGTAACATCTACTCTGGAATTAATGTGG 2998
Db 2941 GAATTCATCATCACATTTTGTATACCTGCCATGTAACATCTACTCTGGAATTAATGTGG 3000
Qy TGCCATGAGGGGCTACGGGTTCTGTTAGTAACTCTAAATACCTTAACCCCATCTGCTAA 3058
Db 3001 TGCCATGAGGGGCTACGGGTTCTGTTAGTAACTCTAAATACCTTAACCCCATCTGCTAA 3060
Qy AATCAAAATGATACATATCTCTGAGAGACCCAGCAATACCAATGAAATTAATCAAAAAAAA 3118
Db 3061 AATCAAAATGATACATATCTCTGAGAGACCCAGCAATACCAATGAAATTAATCAAAAAAAA 3120
Qy 3119 AA 3120
Db 3121 AA 3122

RESULT 8
AY354202 3030 bp mRNA linear PRI 02-SEP-2003
LOCUS Homo sapiens dipeptidyl peptidase 8 isoform 3 (DPP8) mRNA, complete cds, alternatively spliced.
ACCESSION AY354202
VERSION Submitted (29-JUL-2003) Key Lab of Reproductive Medicine, Nanjing Medical University, 140 Han Zhong Road, Nanjing, Jiangsu 210029, China
KEYWORDS Location/Qualifiers
SOURCE 1..3030
ORGANISM /organism="Homo sapiens"
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3030)
AUTHORS Sha, J. H., Zhou, Z. M. and Li, J. M.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-2003) Key Lab of Reproductive Medicine, Nanjing Medical University, 140 Han Zhong Road, Nanjing, Jiangsu 210029, China
FEATURES
source
1..3030
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/tissue="testis"
1..3030
/gene="DPP8"
161..2857
/gene="DPP8"
/notes="alternatively spliced"
/codon_start=1
/product="dipeptidyl peptidase 8 isoform 3"
/protein_id="AAQ63887.1"
/db_xref="GI:34329352"
/translation="MWKSEQMKIKSGKCNNAAMETEQLGVEIFETADCEINISQD
RPLLEFFVYRYSWQLKLLADTKYHGYNMAKAPHDFMFKVRNDPDGPHSDRIYVL
```

ORIGIN

```
Query Match 88.3%; Score 2756; DB 5; Length 3030;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2806; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 201 AGAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTGGGTCTTGAGATATTTGA 260
Db 196 AGAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTGGGTCTTGAGATATTTGA 255
Qy 261 AACTCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCCTAAATTTGAGCCTTTTA 320
Db 256 AACTCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCCTAAATTTGAGCCTTTTA 315
Qy 321 TGTTCAGCGGTATTCTCTGGAGTCAAGTAAAGAGTCTTCCGCGATACCAAGAAATATCA 380
Db 316 TGTTCAGCGGTATTCTCTGGAGTCAAGTAAAGAGTCTTCCGCGATACCAAGAAATATCA 375
Qy 381 TGGCTACATGATGCTAAGGCACACATGATTTTCATGTTGTGAAGAGAAATGATCCAGA 440
Db 376 TGGCTACATGATGCTAAGGCACACATGATTTTCATGTTGTGAAGAGAAATGATCCAGA 435
Qy 441 TGGACCTCATTCAGACAGAAATCTATTACCTTGCATCTCTGGTGAACAGAGAAATAC 500
Db 436 TGGACCTCATTCAGACAGAAATCTATTACCTTGCATCTCTGGTGAACAGAGAAATAC 495
Qy 501 ACTGTTTATCTGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTC 560
Db 496 ACTGTTTATCTGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTC 555
Qy 561 GAAGCTCTTTTGGATCTTTTTCAGGCAACATCTGGAATGGAATGATTTCTCGAGAAGA 620
Db 556 GAAGCTCTTTTGGATCTTTTTCAGGCAACATCTGGAATGGAATGATTTCTCGAGAAGA 615
Qy 621 AGAATATTAAAGAGAAAGAAACGCAATGGAACAGTTCGGAATTCGTTTACGATTATCA 680
Db 616 AGAATATTAAAGAGAGAAAGAAACGCAATGGAACAGTTCGGAATTCGTTTACGATTATCA 675
Qy 681 CCAAGGAAGTGAACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGATGG 740
Db 676 CCAAGGAAGTGAACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGATGG 735
Qy 741 AGGCGCCACAAGGATTTTACGCAACATCTTAAAGGCCCATCTAGTGGAAATCTAGTTGCC 800
Db 736 AGGCGCCACAAGGATTTTACGCAACATCTTAAAGGCCCATCTAGTGGAAATCTAGTTGCC 795
Qy 801 CAACATACGATGATTCGCAAAATTTATGCCCGCTGATCCAGATCGATGCTTTTATACA 860
Db 796 CAACATACGATGATTCGCAAAATTTATGCCCGCTGATCCAGATCGATGCTTTTATACA 855
Qy 861 TAGCAACGATATTTGGATATCTAAATCGTACACAGAGAGAAAGGAGACTCATTATGT 920
Db 856 TAGCAACGATATTTGGATATCTAAATCGTACACAGAGAGAGGAGACTCATTATGT 915
Qy 921 GCACATGAGCTAGCACAACATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 980
Db 916 GCACATGAGCTAGCACAACATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 975
Qy 981 TCTCCAGAGAGATTTGATAGATATTTGGGCTATTGGGTGGTGTCCAAAGCTGAACAAC 1040
```

```
AMSGENRNTLFYSEIPKTIINRAAVLMSWKLDDLEQATLDYCMVSEBELRERKR
IGTVGASVDYHQSSTFLFOAGSGIVHKDGGQGGTQQLRNLVETSCPNLRMDP
KLCPADPWIAFIHSNDLIWISNITREBRRLTYVNBLANMEDEASAGVATFVLQBE
FRYSGYMWCPKAETTPSGGKILRIIYEENDESVEIHIWTSPLMSTRADSRYPKPT
GTANPKVTFKMSIEMIDAGRIIDVIXKELQFSEILFEGVEYIARAGWTPEGRYKPT
ILLDRSOTLOIIVLISPLFIPIVEDDUNWRONRLIESVPDSVPLIIYEETDIIWINIH
DIPHPQPSHREBEIPIFASCKTGFRHLKIITSLKSKRSGGLPAPSPFKCPI
KEEIAITSGEWEVLGRSGNLOVDERRRLVTFEGTKDSPLKHLHYVSYVNPVEVRL
TDRGYSHCCCTSOHCDFISATSNQKPHCVSLKSLSPDEDDPTCKTKEFWATILISA
GRLPYTPREIFESFETITGLIONLYKHLQKPKYPTVLFTGGPQVOLVNNRPFK
GVKYRELLASLGVVYVILNDRGSCHRGLKFEAGFYKMGQIEIDVDEGLQYLASR
VAFIDDRVHGWHSYVILSNMALMORSIDIFVAIAGAPYTLWIFDDGTGTYRGMGH
PQNSQGGIILGSAVMAEKFPFENRLLHGLFDLENFHAFTSILLSLFLVRAGKPYD
LQIYFGRHSIRVPESGHEHYELHLLHLOENLGSRIAALKVI"
```

Db 976 TCTCCAGAGAAATTTGATAGATATCTGGCTATTTGGTGTCCAAAGCTGAAACAAC 1035
 QY 1041 TCCAGATGGTGGTAAATTTCTTAGAATTTCTATATGAAGAAATGATGATCTGAGGTGGA 1100
 Db 1036 TCCAGATGGTGGTAAATTTCTTAGAATTTCTATATGAAGAAATGATGATCTGAGGTGGA 1095
 QY 1101 AATTATTCATGCTTACATCCCTATGTTGGAAACAAAGAGGCGAGATTCATTCGGTTATCC 1160
 Db 1096 AATTATTCATGCTTACATCCCTATGTTGGAAACAAAGAGGCGAGATTCATTCGGTTATCC 1155
 QY 1161 TAAACAGGATACAGAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATGATGATGTC 1220
 Db 1156 TAAACAGGATACAGAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATGATGATGTC 1215
 QY 1221 TGAAGGAGGATCATAGATGTCATAGATAAGGAATTAATTTCAACCTTTTGGATTCATTT 1280
 Db 1216 TGAAGGAGGATCATAGATGTCATAGATAAGGAATTAATTTCAACCTTTTGGATTCATTT 1275
 QY 1281 TGAAGGAGGATGAAATATATGTCAGAGCTGGATGGACTCTCTGAGGGAAATATGCTTGGTC 1340
 Db 1276 TGAAGGAGGATGAAATATATGTCAGAGCTGGATGGACTCTCTGAGGGAAATATGCTTGGTC 1335
 QY 1341 CATCTACTAGATCGCTCCAGACTCCCTACAGATAGTGTGATCTCACCTGAATTAATTT 1400
 Db 1336 CATCTACTAGATCGCTCCAGACTCCCTACAGATAGTGTGATCTCACCTGAATTAATTT 1395
 QY 1401 TATCCAGGTAGAAGATGATGTTATGGAAGGCGAGAGACTCATTTAGTCACTGCTGATTC 1460
 Db 1396 TATCCAGGTAGAAGATGATGTTATGGAAGGCGAGAGACTCATTTAGTCACTGCTGATTC 1455
 QY 1461 TGTGAGCGCACTAATTAATCTATGAAGAAACAAAGAGATCTGATTAATATCATGACAT 1520
 Db 1456 TGTGAGCGCACTAATTAATCTATGAAGAAACAAAGAGATCTGATTAATATCATGACAT 1515
 QY 1521 CTTTCACTGTTTTCCCAAGTCCAGAGAGGAAATTTAGTGTATTTTGGCTCTGAAATG 1580
 Db 1516 CTTTCACTGTTTTCCCAAGTCCAGAGAGGAAATTTAGTGTATTTTGGCTCTGAAATG 1575
 QY 1581 CAAAAAGGTTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAA 1640
 Db 1576 CAAAACAGGTTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAA 1635
 QY 1641 AGATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGTCCTATCAAGAGAGAT 1700
 Db 1636 AGATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGTCCTATCAAGAGAGAT 1695
 QY 1701 AGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGA 1760
 Db 1696 AGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGA 1755
 QY 1761 TGAAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAGATCTCCCTTTAGAGCATCACCT 1820
 Db 1756 TGAAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAGATCTCCCTTTAGAGCATCACCT 1815
 QY 1821 GTACCTGATGATGATGATTAATCTTGGAGAGTGCACAAAGCTGACCTGGCTGCTACTC 1880
 Db 1816 GTACCTGATGATGATGATTAATCTTGGAGAGTGCACAAAGCTGACCTGGCTGCTACTC 1875
 QY 1881 ACATTTCTGCTGATCAGTCAGCACTGTGATCTCTTTTAAAGTAAAGTATAGTAACAGAA 1940
 Db 1876 ACATTTCTGCTGATCAGTCAGCACTGTGATCTCTTTTAAAGTAAAGTATAGTAACAGAA 1935
 QY 1941 GAATCCACTGTGTGCTTCTTACAGCTATCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAG 2000
 Db 1936 GAATCCACTGTGTGCTTCTTACAGCTATCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAG 1995
 QY 2001 AACAAAGGAATTTTGGGCCACCATTTTGGATTTCAAGAGTCTCTCTTCACTATATCTCC 2060
 Db 1996 AACAAAGGAATTTTGGGCCACCATTTTGGATTTCAAGAGTCTCTCTTCACTATATCTCC 2055
 QY 2061 TCCAGAAATTTTCTCTTTTGAAGTATCTAGGATTTATGATGAGGATGCTCTACAA 2120

Db 2056 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAA 2115
 QY 2121 GCCTCATGATCTACAGCTCGAAAGAAATATCTTACTGTCTGTCTTATATATATGCTGGTCC 2180
 Db 2116 GCCTCATGATCTACAGCTCGAAAGAAATATCTTACTGTCTGTCTTATATATATGCTGGTCC 2175
 QY 2181 TCAGGTGAGTGTGGTGAATTAATTCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTT 2240
 Db 2176 TCAGGTGAGTGTGGTGAATTAATTCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTT 2235
 QY 2241 AGCCTCTCTAGGTTATGTTGTAGTAGATAGACAAACAGGGGATCTCTGTCCACGAGGCT 2300
 Db 2236 AGCCTCTCTAGGTTATGTTGTAGTAGATAGACAAACAGGGGATCTCTGTCCACGAGGCT 2295
 QY 2301 TAAATTTGAAGCGCTTTTAAATATAAAATGGGTCAAATAGAAATTCAGCATCAGGTGGA 2360
 Db 2296 TAAATTTGAAGCGCTTTTAAATATAAAATGGGTCAAATAGAAATTCAGCATCAGGTGGA 2355
 QY 2361 AGGATCCCAATATCTAGCTTCTCGATATGATTTCAATTCAGTTCGTGGGATCTTCA 2420
 Db 2356 AGGATCCCAATATCTAGCTTCTCGATATGATTTCAATTCAGTTCGTGGGATCTTCA 2415
 QY 2421 CGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTTAATGACAGAGTCAAGATATCTT 2480
 Db 2416 CGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTTAATGACAGAGTCAAGATATCTT 2475
 QY 2481 CAGGCTGGTCTATGCTGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAC 2540
 Db 2476 CAGGCTGGTCTATGCTGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAC 2535
 QY 2541 GGAACCTTATATGGGTCAACCTGACCAAGTGAACAGGGCTATTTACCTTAGGATCTGTGGC 2600
 Db 2536 GGAACCTTATATGGGTCAACCTGACCAAGTGAACAGGGCTATTTACCTTAGGATCTGTGGC 2595
 QY 2601 CATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTCTTCAATGTTTCTT 2660
 Db 2596 CATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTCTTCAATGTTTCTT 2655
 QY 2661 GGNATGATGTCGATTTTGGACATACAGTATATTAATGAGTGTGTGTAGTGGGCTGG 2720
 Db 2656 GGNATGATGTCGATTTTGGACATACAGTATATTAATGAGTGTGTGTAGTGGGCTGG 2715
 QY 2721 AAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAGAGTTCTCTGAATC 2780
 Db 2716 AAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAGAGTTCTCTGAATC 2775
 QY 2781 GGAGAACATTAATGATGATCTGTCATCTTTTGGCATCTA CTTTCAAGAAAACTTGGATCACGTAT 2840
 Db 2776 GGAGAACATTAATGATGATCTGTCATCTTTTGGCATCTA CTTTCAAGAAAACTTGGATCACGTAT 2835
 QY 2841 TGCTGCTCTAAGATGATATATTTTGGACCTGTGTAGAACTCTCTGGGTATACACCTGGCTA 2900
 Db 2836 TGCTGCTCTAAGATGATATATTTTGGACCTGTGTAGAACTCTCTGGGTATACACCTGGCTA 2895
 QY 2901 TTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATGATTCATCATTTTGAAT 2960
 Db 2896 TTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATGATTCATCATTTTGAAT 2955
 QY 2961 ACCTGGCATGATACATCTACTCTCTGAAATTAATGCTGGTGCATGCA 3007
 Db 2956 ACCTGGCATGATACATCTACTCTCTGAAATTAATGCTGGTGCATGCA 3002

RESULT 9
 AX354795
 LOCUS AX354795 2649 bp DNA linear PAT 06-FEB-2002
 DEFINITION Sequence 3 from Patent WO0179473.
 ACCESSION AX354795
 VERSION AX354795.1 GI:18619528
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Submitted (04-NOV-2002) Ferring Research Institute, 3550 General Atomics Ct., San Diego, CA 92121, USA		JOURNAL	
Location/Qualifiers		FEATURES	
1..2656		source	
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosome="15"			
/map="15q22.1-q22.2"			
1..2656		gene	
/gene="DPRP1"			
8..2656		CDS	
/gene="DPRP1"			
/codon_start=1			
/product="dipeptidyl peptidase IV-related protein-1"			
/protein_id="AA017261.1"			
/db_xref="GI:27549550"			
/translation="MAAMETEQLGVEIFETADCEENIESQDRPKLEPFVVERYSWSQ LKKJLADTRKYHGMKAKAFDHFVVRNDPDPGHSRIYIYLAMSGENRENTLYSEI PKTINRAAVLMSWKPLDLDFQATLDYGMYSREBELRERKRIGTGVGASIDYHQSG TFLFOAGSGIVHVGQGGPQGTQOPLRPNLVETSCPNRMDPKLCPADPDWIAPIHSN DIWISNIVTRERRLTYYVHNLAMBEDARSAGVATFVLOEEDFYSGYVWYCPKAEIT PSQKILRILYEENDESEVEIHHVTSPLMETRRADSPRYKPTGTANPKVTFKMSIEMI DAEGRLIDVIDKELLQPELIFEGVEYLARAGWTPGKYAWSILLDRSQRLQIULIS PELFIPVEDDVMERQRLIESVPSVTPLIIEETTDIWINIHDIHVFPQSHREELER IFASECKTGFRLHYKITISILKESKYKSSGGLPAPSPDFKPIKEEIAITSGEWEVLGR HGSNIQVDRERLVYFEGTKDSPLEHLYVYVYVNGEVTRLDTRGYSHSCCISQHCDD TPTSKYSNOKPHCVSLKLSPPEDDPTCKTKEFWATILDSAGPLPYDTPPTPFSPES TTGFTLGMKYKPHDLOPKKYPTVLFYIGGPOVLVNNRPFKGVYFRLANTLASLGVV VVVIDNRGSGHRLKPEGAFKPKWGOIIEIDDOVSGLOYLASRYDFILDRLVGHGWSY GGYSLSJMLUMQRSDIFRVAIAGAPVTLMIPIYDTGTTRYRMGHGPDQNGQYVVGSAVMO AEKFPSEPNRLLLHFLDNEHFHFAVTSILLSFLVRAGKPYDLQIYQERHSIRVPES GEHYELHLHLYLQENLGSRIALAKVI"			
		ORIGIN	
		Query Match 80.0%; Score 2495; DB 5; Length 2656;	
		Best Local Similarity 99.3%; Pred. No. 0;	
		Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY	213	CATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTGGGACTG	272
Db	7	CATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTGGGACTG	66
QY	273	TGAGAGAAATATGGAATCAGAGATCGGCTAAATTTGGAGCCCTTTTATGTTGACGGTA	332
Db	67	TGAGAGAAATATGGAATCAGAGATCGGCTAAATTTGGAGCCCTTTTATGTTGACGGTA	126
QY	333	TTCTCTGGAGTCAGCTTAAAGCTGCTTGCAGATACAGAAATATCATGGCTACATGAT	392
Db	127	TTCTCTGGAGTCAGCTTAAAGCTGCTTGCAGATACAGAAATATCATGGCTACATGAT	186
QY	393	GGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGATGATCCAGATGGACCTCATTC	452
Db	187	GGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGATGATCCAGATGGACCTCATTC	246
QY	453	AGACAGAAATCTATTACCTTGCATGCTGCTGGTGAAGACAGAGAAATATACACTGTTTATTC	512
Db	247	AGACAGAAATCTATTACCTTGCATGCTGCTGGTGAAGACAGAGAAATATACACTGTTTATTC	306
QY	513	TGAATATCCCAAACTATCAATAGACAGACAGCTTAAATGCTCTCTTGAAGCCTCTTTT	572
Db	307	TGAATATCCCAAACTATCAATAGACAGACAGCTTAAATGCTCTCTTGAAGCCTCTTTT	366
QY	573	GGATCTTTTTCAGCAACACTGGACTATGGAATGATTTCTTCGAGAGAGAACTATTAAAG	632
Db	367	GGATCTTTTTCAGCAACACTGGACTATGGAATGATTTCTTCGAGAGAGAACTATTAAAG	426
QY	633	AGAAAGAAACCGATTGGAAACAGTCCGGAATTCCTTTACGATTATCACCAAGAAAGTGG	692
Db	427	AGAAAGAAACCGATTGGAAACAGTCCGGAATTCCTTTACGATTATCACCAAGAAAGTGG	486
QY	693	AACATTTCTGTTTCAAGCGCGGTAGTGGAAATTTATCATGTAAGATGAGGGCCACAGG	752

Db	487	AACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGCTAAAGATGGAGGCCACAAGG	546
Qy	753	ATTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAATCTAGTGTCCCAATACGGAT	812
Db	547	ATTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAATCTAGTGTCCCAATACGGAT	606
Qy	813	GGATCCAAAATTTATGCCCGCTGATCCAGACTCGGATGCTTTTATACATAGCAAGATAT	872
Db	607	GGATCCAAAATTTATGCCCGCTGATCCAGACTCGGATGCTTTTATACATAGCAAGATAT	666
Qy	873	TTGGATATCTAACATCTTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	932
Db	667	TTGGATATCTAACATCTTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	726
Qy	933	AGCCAAATGGAAG	992
Db	727	AGCCAAATGGAAG	786
Qy	993	ATTTGATAGATATTTCTGGCTATTTGGTGGTGCCTCAAGAGAGAGAGAGAGAGAGAG	1052
Db	787	ATTTGATAGATATTTCTGGCTATTTGGTGGTGCCTCAAGAGAGAGAGAGAGAGAGAG	846
Qy	1053	TAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGT	1112
Db	847	TAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGT	906
Qy	1113	TACATCCCTATGTTGGAAACAAGGAGGCGAGATTCATTCGGTATATCTTAAACAAGGTAC	1172
Db	907	TACATCCCTATGTTGGAAACAAGGAGGCGAGATTCATTCGGTATATCTTAAACAAGGTAC	966
Qy	1173	AGCAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATATGATGATGCTGAAGGAAGAT	1232
Db	967	AGCAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATATGATGATGCTGAAGGAAGAT	1026
Qy	1233	CATAGATGTCATAGATAAGGAATTAATTCACCTTTTGAAGATCTTATTTGAAGGAGTTGA	1292
Db	1027	CATAGATGTCATAGATAAGGAATTAATTCACCTTTTGAAGATCTTATTTGAAGGAGTTGA	1086
Qy	1293	ATATATTTGCGAGAGCTGGATGAGCTTCTGAGGGAATATGCTTGGTCCATCTCTACTAGA	1352
Db	1087	ATATATTTGCGAGAGCTGGATGAGCTTCTGAGGGAATATGCTTGGTCCATCTCTACTAGA	1146
Qy	1353	TCGCTCCAGACTCGCTACAGATAGTGTGATCTCAGCTGAAATATGCTTGAATTTTGAATTT	1412
Db	1147	TCGCTCCAGACTCGCTACAGATAGTGTGATCTCAGCTGAAATATGCTTGAATTTTGAATTT	1206
Qy	1413	AGATGATGTTATGGAAGGCGAGAGACTCAATGAGTCAAGTCCCTGATCTGTGACGCACT	1472
Db	1207	AGATGATGTTATGGAAGGCGAGAGACTCAATGAGTCAAGTCCCTGATCTGTGACGCACT	1266
Qy	1473	AATTTATCTAGAGAAACAACAGACATCTGGATTAATATCCATGACATCTTTTATGTTTTT	1532
Db	1267	AATTTATCTAGAGAAACAACAGACATCTGGATTAATATCCATGACATCTTTTATGTTTTT	1326
Qy	1533	TCGCCAAGTCAAGAGAGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAAACAGGTTT	1592
Db	1327	TCGCCAAGTCAAGAGAGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAAACAGGTTT	1386
Qy	1593	CGGTCAATTTATCAAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGCATCCAGTGG	1652
Db	1387	CGGTCAATTTATCAAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGCATCCAGTGG	1446
Qy	1653	TGGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTATCAAAAGAGAGATAGCAATTTACCAG	1712
Db	1447	TGGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTATCAAAAGAGAGATAGCAATTTACCAG	1506
Qy	1713	TGGTGAATGGGAAGTTCTTGGCCGCGCATGGAATCTTAATATCCAAAGTTTGAAGATCAGAG	1772
Db	1507	TGGTGAATGGGAAGTTCTTGGCCGCGCATGGAATCTTAATATCCAAAGTTTGAAGATCAGAG	1566
Qy	1773	GCTGGTATATTTTGAAGGACCAAGAGACTCCCTTTTAGAGCATCACCTGTCAGTAGTCAG	1832
Db	1567	GCTGGTATATTTTGAAGGACCAAGAGACTCCCTTTTAGAGCATCACCTGTCAGTAGTCAG	1626
Qy	1833	TTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCATTCTTTGCTG	1892
Db	1627	TTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCATTCTTTGCTG	1686
Qy	1893	CATCAGTCAGCAGCTGTGACTTCTTTTATAAAGTAGTATAGTAAACAGAGAATCCACACTG	1952
Db	1687	CATCAGTCAGCAGCTGTGACTTCTTTTATAAAGTAGTATAGTAAACAGAGAATCCACACTG	1746
Qy	1953	TGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCAAAACAAGGAAT	2012
Db	1747	TGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCAAAACAAGGAAT	1806
Qy	2013	TTGGGCCACCATTTTGGATTACAGAGTCTCTCTCTGACTATACCTCTCCAGAAATTTT	2072
Db	1807	TTGGGCCACCATTTTGGATTACAGAGTCTCTCTCTGACTATACCTCTCCAGAAATTTT	1866
Qy	2073	CTCTTTTGAAGTAGTACTACTGANTTTACATTTGATGGATGCTCTACAGCCTCATGATCT	2132
Db	1867	CTCTTTTGAAGTAGTACTACTGANTTTACATTTGATGGATGCTCTACAGCCTCATGATCT	1926
Qy	2133	ACAGCCTGGAAGAAATATCTTACTGCTGCTTTTATATATATGTTGGTCTCAGGTGAGTT	2192
Db	1927	ACAGCCTGGAAGAAATATCTTACTGCTGCTTTTATATATATGTTGGTCTCAGGTGAGTT	1986
Qy	2193	GGTGAATANTCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTTAGG	2252
Db	1987	GGTGAATANTCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTTAGG	2046
Qy	2253	TTATGTGGTGTAGTAGACAAAGGGGATCTGTGTACCCAGGGCTTTAAATTTGAAAGG	2312
Db	2047	TTATGTGGTGTAGTAGACAAAGGGGATCTGTGTACCCAGGGCTTTAAATTTGAAAGG	2106
Qy	2313	CGCCTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATA	2372
Db	2107	CGCCTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATA	2166
Qy	2373	TCTAGCTTCTCGATATGATTTTCAATTTGACTTAGATCGTGTGGGCAATCCAGGCTGTCTTA	2432
Db	2167	TCTAGCTTCTCGATATGATTTTCAATTTGACTTAGATCGTGTGGGCAATCCAGGCTGTCTTA	2226
Qy	2433	TGGAGGATACCTCTCCCTGATGGCAATTAATGACAGAGGTCAAGATATCTTTGAGGTTGCTAT	2492
Db	2227	TGGAGGATACCTCTCCCTGATGGCAATTAATGACAGAGGTCAAGATATCTTTGAGGTTGCTAT	2286
Qy	2493	TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAACTTATAT	2552
Db	2287	TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAACTTATAT	2346
Qy	2553	GGGTCAACCTTGACCAAGATGAAACAGGCTTATTACTTAGGATCTGTGGCCATGCAAGCAGA	2612
Db	2347	GGGTCAACCTTGACCAAGATGAAACAGGCTTATTACTTAGGATCTGTGGCCATGCAAGCAGA	2406
Qy	2613	AAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATCGTTTCTGGATGAGATGT	2672
Db	2407	AAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATCGTTTCTGGATGAGATGT	2466
Qy	2673	CCATTTTGCACATACCAAGTATATTTACTGAGTTTATGAGGCTGGAAGCCATATGA	2732
Db	2467	CCATTTTGCACATACCAAGTATATTTACTGAGTTTATGAGGCTGGAAGCCATATGA	2526
Qy	2733	TTTACAGATCTATCTCAGGAGAGACACAGCATTAAGTCTCTGAATCTGGGAGAACATTA	2792
Db	2527	TTTACAGATCTATCTCAGGAGAGACACAGCATTAAGTCTCTGAATCTGGGAGAACATTA	2586
Qy	2793	TGNACTGATCTTTTGGACTACTCTTCAAGAAAACCTTGGATACGATATGCTGCTTAAA	2852
Db	2587	TGNACTGATCTTTTGGACTACTCTTCAAGAAAACCTTGGATACGATATGCTGCTTAAA	2646
Qy	2853	AGTGATAT	2860
Db	2647	AGTGATAT	2654

RESULT 11	AR631278	2671 bp	DNA	linear	PAT 14-FEB-2005
LOCUS	Sequence 2 from patent US 6844180.				
DEFINITION	AR631278				
ACCESSION	AR631278.1	GI:59770920			
VERSION	Unknown.				
KEYWORDS	Unclassified.				
SOURCE	Ol,S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.				
ORGANISM	Serine protease genes related to DPPIV				
REFERENCE	Patent: US 6844180-A 2 18-JAN-2005;				
AUTHORS	Perring BV,;				
TITLE	NLX;				
JOURNAL					
FEATURES	Location/Qualifiers				
source	1..2671				
ORIGIN	/organism="unknown"				
	/mol_type="genomic DNA"				
Query Match	80.0%;	Score 2495;	DB 2;	Length 2671;	
Best Local Similarity	99.9%;	Pred. No. 0;			
Matches 2643;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;	
Qy	213	CATGGCAGCAGCAATGGAAACAGAACAGACAGCTGGGGTTCGAGATATTTGAAACTGGGACTG	272		
Db	7	CATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTTGAGATATTTGAAACTGGGACTG	66		
Qy	273	TGAGGGAATTTGAAATCACAGATCGGCTAAATTTGGAGCTTTTATGTTGAGCGGTA	332		
Db	67	TGAGGGAATTTGAAATCACAGATCGGCTAAATTTGGAGCTTTTATGTTGAGCGGTA	126		
Qy	333	TTCTCGAGTCAGCTTAAAAAGCTGCTCCGATACACAGAAATATCATGCTCATCATGAT	392		
Db	127	TTCTCGAGTCAGCTTAAAAAGCTGCTCCGATACACAGAAATATCATGCTCATCATGAT	186		
Qy	393	GGCTAAGCACCACATGATTTTCATGTTTGTGAAGAGGATGATCCAGATGACCTCATTC	452		
Db	187	GGCTAAGCACCACATGATTTTCATGTTTGTGAAGAGGATGATCCAGATGACCTCATTC	246		
Qy	453	AGACAGAAATCTATTACCTTGCCATGTCGTGTGAGAACAGAGAAATAACCTGTTTATTTC	512		
Db	247	AGACAGAAATCTATTACCTTGCCATGTCGTGTGAGAACAGAGAAATAACCTGTTTATTTC	306		
Qy	513	TGAAATTTCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAAGCCTCTTTT	572		
Db	307	TGAAATTTCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAAGCCTCTTTT	366		
Qy	573	GGATCTTTTTCAGGCAACACTGGACTATGGAATTTATCATCGATGGAAGAACTATTAAAG	632		
Db	367	GGATCTTTTTCAGGCAACACTGGACTATGGAATTTATCATCGATGGAAGAACTATTAAAG	426		
Qy	633	AGAAAGAAACCGAATTGGACAGTCCGAATTCCTTACAGATATCACCAAGGAAGTGG	692		
Db	427	AGAAAGAAACCGAATTGGACAGTCCGAATTCCTTACAGATATCACCAAGGAAGTGG	486		
Qy	693	AACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCATCGATGGAAGAACTATTAAAG	752		
Db	487	AACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCATCGATGGAAGAACTATTAAAG	546		
Qy	753	ATTTACGCAACAACTTTAAGGCCAATCTAGTGGAACTAGTGTGCTCCCAACATACGGAT	812		
Db	547	ATTTACGCAACAACTTTAAGGCCAATCTAGTGGAACTAGTGTGCTCCCAACATACGGAT	606		
Qy	813	GGATCCAAAATTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT	872		
Db	607	GGATCCAAAATTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT	666		
Qy	873	TTGGATATCTTAACATCGTAACAGAGAAAGGAGCTCACTTATGTGCACATGAGCT	932		

Qy	2013	TTGGGCCACATTTTGGATTTACAGAGGTCCTCTTCTGACTACTACTCTCCAGAAATTTT	2072
Db	1807	TTGGGCCACATTTTGGATTTACAGAGGTCCTCTTCTGACTACTACTCTCCAGAAATTTT	1866
Qy	2073	CTCTTTTGAAGTACTACTTGGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCT	2132
Db	1867	CTCTTTTGAAGTACTACTTGGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCT	1926
Qy	2133	ACAGCCTGGAAGAAGAAATATCCTTACTGTGCTGTTCAATATATGGTGGTCTCTAGGTGCAGTT	2192
Db	1927	ACAGCCTGGAAGAAGAAATATCCTTACTGTGCTGTTCAATATATGGTGGTCTCTAGGTGCAGTT	1986
Qy	2193	GGTGAATATCCGTTTAAAGGAGTCAGATTTTCCGTTTGAATACCTTAGCCTCTCTAGG	2252
Db	1987	GGTGAATATCCATTTTAAAGGAGTCAGATTTTCCGTTTGAATACCTTAGCCTCTCTAGG	2046
Qy	2253	TTATGGGTGTAGTGATAGACAACAGGGGATCCTGTCAACGGAGGCTTAAATTTGAAGG	2312
Db	2047	TTATGGGTGTAGTGATAGACAACAGGGGATCCTGTCAACGGAGGCTTAAATTTGAAGG	2106
Qy	2313	CGCCTTTAAATATAAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAATA	2372
Db	2107	CGCCTTTAAATATAAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAATA	2166
Qy	2373	TCTAGCTTCTCGATATGATTTCAITGACTAGATCGTGTGGCATCACGGCTGGTCCCTA	2432
Db	2167	TCTAGCTTCTCGATATGATTTCAITGACTAGATCGTGTGGCATCACGGCTGGTCCCTA	2226
Qy	2433	TGGAGGATACCTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTCAGGGTTGCTAT	2492
Db	2227	TGGAGGATACCTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTCAGGGTTGCTAT	2286
Qy	2493	TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACGTTTATAT	2552
Db	2287	TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACGTTTATAT	2346
Qy	2553	GGGTCAACCTCACCAGAAATGAACGGGCTATTACTAGATCTGTGGCCATGCAAGCAGA	2612
Db	2347	GGGTCAACCTCACCAGAAATGAACGGGCTATTACTAGATCTGTGGCCATGCAAGCAGA	2406
Qy	2613	AAAGTTCCTCTGAAACCAATCGTTTACTGTCTTACATGTTTCTTGGATGAGAATGT	2672
Db	2407	AAAGTTCCTCTGAAACCAATCGTTTACTGTCTTACATGTTTCTTGGATGAGAATGT	2466
Qy	2673	CCATTTTGGCATACACAGTATATATCTAGTGTTTTTTGTAGGGGTGGAAGCCCATATGA	2732
Db	2467	CCATTTTGGCATACACAGTATATATCTAGTGTTTTTTGTAGGGGTGGAAGCCCATATGA	2526
Qy	2733	TTTACAGATCTATCCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAGCAATTA	2792
Db	2527	TTTACAGATCTATCCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAGCAATTA	2586
Qy	2793	TGAATCGCATCTTTTGCATCTACCTTCAAGAAAACTTGGATCACGTATTTGCTCTTAAA	2852
Db	2587	TGAATCGCATCTTTTGCATCTACCTTCAAGAAAACTTGGATCACGTATTTGCTCTTAAA	2646
Qy	2853	AGTGATAT 2860	
Db	2647	AGTGATAT 2654	

[illegible]

Qy	1053	TAAAAATCTTTAGAAATTC	TATATGAAGAAAAATG	AAATCTGAGGTG	AAAAATTA	TATTCATGT	1112
Db	847	TAAAAATCTTTAGAAATTC	TATATGAAGAAAAATG	AAATCTGAGGTG	AAAAATTA	TATTCATGT	906
Qy	1113	TACATCCCCTATGTTGG	AAAAAACAAGAGGGC	AGATTCATTCGGTTAT	CCTAAAAACAGGTAC	1172	
Db	907	TACATCCCCTATGTTGG	AAAAAACAAGAGGGC	AGATTCATTCGGTTAT	CCTAAAAACAGGTAC	966	
Qy	1173	AGCAAAATCCTAAAGT	CACATTTTAAAGATGT	CAGAAATAATGAT	TGATCTGAAGAAAGGAT	1232	
Db	967	AGCAAAATCCTAAAGT	CACATTTTAAAGTGT	CAGAAATAATGAT	TGATCTGAAGAAAGGAT	1026	
Qy	1233	CATAGATGTCATAGATA	AGGAACTAAATTCAC	CCCTTTGAGATTC	TATTTGAGGAGTTGA	1292	
Db	1027	CATAGATGTCATAGATA	AGGAACTAAATTCAC	CCCTTTGAGATTC	TATTTGAGGAGTTGA	1086	
Qy	1293	ATATATTTGCCAGAGCT	GGATGGACTCTCTGAGG	AAAAATATGCTTGGT	CCATCTCTACTAGA	1352	
Db	1087	ATATATTTGCCAGAGCT	GGATGGACTCTCTGAGG	AAAAATATGCTTGGT	CCATCTCTACTAGA	1146	
Qy	1353	TCGCTCCCGAGCTCG	CTACAGATAGTGT	TGATCTCACCTG	AAATTAATTAATTC	CCAGTAGA	1412
Db	1147	TCGCTCCCGAGCTCG	CTACAGATAGTGT	TGATCTCACCTG	AAATTAATTAATTC	CCAGTAGA	1206
Qy	1413	AGATGATGTTATG	AAAAAGCAGAGACT	CAATGAGTCAGT	GCCTGATTTCTGTG	ACGCCACT	1472
Db	1207	AGATGATGTTATG	AAAAAGCAGAGACT	CAATGAGTCAGT	GCCTGATTTCTGTG	ACGCCACT	1266
Qy	1473	AAATATCTATGA	AAAAAACAACAGACAT	CTGGATAAATAT	TCCATGACATCTTT	CATGTTTT	1532
Db	1267	AAATATCTATGA	AAAAAACAACAGACAT	CTGGATAAATAT	TCCATGACATCTTT	CATGTTTT	1326
Qy	1533	TCCCCAAAGTC	ACGAGGAAATTG	AGTTATTTTTGCTCT	GATGCAAAACAGGTTT	1592	
Db	1327	TCCCCAAAGTC	ACGAGGAAATTG	AGTTATTTTTGCTCT	GATGCAAAACAGGTTT	1386	
Qy	1593	CGGTCAATTTAT	CAAAAATTA	CATCTATTTTTAAAGG	AAAGCAAAATATAA	ACGATCCAGTGG	1652
Db	1387	CGGTCAATTTAT	CAAAAATTA	CATCTATTTTTAAAGG	AAAGCAAAATATAA	ACGATCCAGTGG	1446
Qy	1653	TGGGCTGCCTGCT	CCAAAGTAAATTTCA	AGTGCCTATCA	AAAGAGGAGATAG	CAATTAACCG	1712
Db	1447	TGGGCTGCCTGCT	CCAAAGTAAATTTCA	AGTGCCTATCA	AAAGAGGAGATAG	CAATTAACCG	1506
Qy	1713	TGGTGAATGGGAAGT	CTTTGGCCGGCAT	TGGATCTAATATCCA	AGTTGATGAAGT	CAGAAAG	1772
Db	1507	TGGTGAATGGGAAGT	CTTTGGCCGGCAT	TGGATCTAATATCCA	AGTTGATGAAGT	CAGAAAG	1566
Qy	1773	GCTGGTATATTTT	TGAAGGCACCAAG	AGCTCCCTCTTTAG	AGCATCACCTGT	TACGTAGTACG	1832
Db	1567	GCTGGTATATTTT	TGAAGGCACCAAG	AGCTCCCTCTTTAG	AGCATCACCTGT	TACGTAGTACG	1626
Qy	1833	TTACGTAAATCT	GTGAGAGGTG	CAAGGCTGACTG	ACCGTGGCTACT	CAATTTCTTGCTG	1892
Db	1627	TTACGTAAATCT	GTGAGAGGTG	CAAGGCTGACTG	ACCGTGGCTACT	CAATTTCTTGCTG	1686
Qy	1893	CATCAGTCAGCACT	GTGACTCTTTTAT	AGTAAGTATAGTA	TAAACAGAGAAAT	CCACACTG	1952
Db	1687	CATCAGTCAGCACT	GTGACTCTTTTAT	AGTAAGTATAGTA	TAAACAGAGAAAT	CCACACTG	1746
Qy	1953	TGTGTCCCTTTT	TACAGACTAT	CAAGTCTG	TGAAGTACCAACTTTG	CAAAAACAAAGGAAT	2012
Db	1747	TGTGTCCCTTTT	TACAGACTAT	CAAGTCTG	TGAAGTACCAACTTTG	CAAAAACAAAGGAAT	1806
Qy	2013	TTGGGCCACCATTT	TTGGATTCAGC	AGGTCCTCTTCTG	ACTATACTCTCCAG	AAAAATTTT	2072
Db	1807	TTGGGCCACCATTT	TTGGATTCAGC	AGGTCCTCTTCTG	ACTATACTCTCCAG	AAAAATTTT	1866
Qy	2073	CTCTTTTGA	AAAGTACTACTG	GAATTAATG	TGATGGGATG	CTCTACAGCCCTCAT	2132
Db	1867	CTCTTTTGA	AAAGTACTACTG	GAATTAATG	TGATGGGATG	CTCTACAGCCCTCAT	1926
Qy	2133	ACAGCCCTG	GAAGAAATAT	CTCTACTGTG	CTGTTTCATATAT	TGTTGCTCTCAGGTC	2192

Db	1927		ACAGCCTGGAAAGAAATACCTCTGCTGCTTCATATATGGTGGTCTCTCAGGTGCAGTT	1986
Qy	2193		GGTGAATATCCGTTTAAAGAGAGTCAAGATATTTCCGCTTGAATACCCCTAGCCCTCTCTAGG	2252
Db	1987		GGTGAATATCCGTTTAAAGAGTCAAGATATTTCCGCTTGAATACCCCTAGCCCTCTCTAGG	2046
Qy	2253		TTATGTGGTTGTAGTAGATAGACAACAGGGGATCTCTGTCACCGAGGGCTTAAATTTTGAAGG	2312
Db	2047		TTATGTGGTTGTAGTAGATAGACAACAGGGGATCTCTGTCACCGAGGGCTTAAATTTTGAAGG	2106
Qy	2313		CGCCTTTAAATATAAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCAATA	2372
Db	2107		CGCCTTTAAATATAAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCAATA	2166
Qy	2373		TCTAGCTTCTCGATATGATTTCAATTGACTTAGATCTGTGGGCATCCAGGCTGGTTCCTA	2432
Db	2167		TCTAGCTTCTCGATATGATTTCAATTGACTTAGATCTGTGGGCATCCAGGCTGGTTCCTA	2226
Qy	2433		TGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTAT	2492
Db	2227		TGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTAT	2286
Qy	2493		TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATAT	2552
Db	2287		TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATAT	2346
Qy	2553		GGGTCAACCTGACAGNATGACAGGGCTATTACTTTAGGATCTCTGGCCATCGAAGCAGA	2612
Db	2347		GGGTCAACCTGACAGNATGACAGGGCTATTACTTTAGGATCTCTGGCCATCGAAGCAGA	2406
Qy	2613		AAAGTTCCCTCTCGAACCAAAATCGTTTACTGCTCTTTCATCGTTTCTCGATGAGATGT	2672
Db	2407		AAAGTTCCCTCTCGAACCAAAATCGTTTACTGCTCTTTCATCGTTTCTCGATGAGATGT	2466
Qy	2673		CAATTTTGGACATACAGTATATATCTGAGTGTTTTTTAGTGAAGGCTGGAAAGCCATATGA	2732
Db	2467		CAATTTTGGACATACAGTATATATCTGAGTGTTTTTTAGTGAAGGCTGGAAAGCCATATGA	2526
Qy	2733		TTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCCTGAACTCGGAGAACATTA	2792
Db	2527		TTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCCTGAACTCGGAGAACATTA	2586
Qy	2793		TGAATCTGCATCTTTTGGACTACTCTTCAAGAAACCTTTGGATCACCTGATTGCTGCTCTAAA	2852
Db	2587		TGAATCTGCATCTTTTGGACTACTCTTCAAGAAACCTTTGGATCACCTGATTGCTGCTCTAAA	2646
Qy	2853		AGTGATAT 2860	
Db	2647		AGTGATAT 2654	
RESULT 13				
AB448400				
LOCUS				
Accession 1103 from patent US 6673549.				
2797 bp DNA linear PAT 20-FEB-2004				

RESULT 13
AR448400
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN

linear PAT 20-FEB-2004
2797 bp DNA
1103 from patent US 6673549.
Sequence
AR448400
AR448400.1 GI:42676724
Unknwn.
Unknwn.
Unclassified.
1 (bases 1 to 2797)
Furness,L.M. and Buchbinder,J.L.
Genes expressed in C3A liver cell cultures treated with steroids
Patent: US 6673549-A 1103 06-JAN-2004;
Inocyte Corporation; Palo Alto, CA
Location/Qualifiers
1. .2797
/organism="unknown"
/mol_type="genomic DNA"

Query Match									
Best Local Similarity 74.2%; Score 2315; DB 2; Length 2797;									
Matches 2785; Conservative 99.8%; Pred. No. 0; Mismatches 1; Indels 5; Gaps 2;									
Qy	335	CTTGGAGTCAGCTTAAAGCTGCTTGGGATACAGAAATATCATGGCTACATGATGG	394						
Db	1	CTTGGAGTCAGCTTAAAGCTGCTTGGGATACAGAAATATCATGGCTACATGATGG	60						
Qy	395	CTAAGGCCACCATCATATTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG	454						
Db	61	CTAAGGCCACCATCATATTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG	120						
Qy	455	ACAGAACTCATTTACCTTGGCCATGCTCTGGTGAGAAACAGAGAAATAACACTGTTTATTCG	514						
Db	121	ACAGAACTCATTTACCTTGGCCATGCTCTGGTGAGAAACAGAGAAATAACACTGTTTATTCG	180						
Qy	515	AAATTCCTCAAAACTATCAATAGACGAGCTTTAATGCTCTCTTGGAAAGCCTCTTTTGG	574						
Db	181	AAATTCCTCAAAACTATCAATAGACGAGCTTTAATGCTCTCTTGGAAAGCCTCTTTTGG	240						
Qy	575	ATCTTTTTCAGGCAACACTGGAATGATTTCTCGAAGAGAAAGAACTATTAAGAG	634						
Db	241	ATCTTTTTCAGGCAACACTGGAATGATTTCTCGAAGAGAAAGAACTATTAAGAG	300						
Qy	635	AAAGAAACCGATTTGGAAACAGTGGAAATGCTTCTTACGATTTACCAAGAGGAAGTGAA	694						
Db	301	AAAGAAACCGATTTGGAAACAGTGGAAATGCTTCTTACGATTTACCAAGAGGAAGTGAA	360						
Qy	695	CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAGAAAGATGGAGGGCCACAAAGAT	754						
Db	361	CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAGAAAGATGGAGGGCCACAAAGAT	420						
Qy	755	TTACGCAACAACTTTAAGGCCCAATCTAGTGMAACTAGTTGTCCCAACATACGGATGG	814						
Db	421	TTACGCAACAACTTTAAGGCCCAATCTAGTGMAACTAGTTGTCCCAACATACGGATGG	480						
Qy	815	ATCCAAATATATGCCCCGCTGATCCAGACTGGATTTGCTTTATACATACAGCAACATATTT	874						
Db	481	ATCCAAATATATGCCCCGCTGATCCAGACTGGATTTGCTTTATACATACAGCAACATATTT	540						
Qy	875	GGATATCTAACTGTAACCCAGAGAAAGAGAGACTCACTATGTGCGACATCAGCTAG	934						
Db	541	GGATATCTAACTGTAACCCAGAGAAAGAGAGACTCACTATGTGCGACATCAGCTAG	600						
Qy	935	CCAACTGGGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAAAT	994						
Db	601	CCAACTGGGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAAAT	660						
Qy	995	TTGATAGATATCTGGCTATTTGGTGTCCAAAGCTGAAAGAACTCCCGCTGGTGGA	1054						
Db	661	TTGATAGATATCTGGCTATTTGGTGTCCAAAGCTGAAAGAACTCCCGCTGGTGGA	720						
Qy	1055	AAATTTCTAGAAATCTATATGAAGAAATGATGATCTGAGGTGGAATTTATTCATGTTA	1114						
Db	721	AAATTTCTAGAAATCTATATGAAGAAATGATGATCTGAGGTGGAATTTATTCATGTTA	780						
Qy	1115	CATCCCTATGTTGGAAACAGG-AGGCGAGATTCATTCGGTTATCCTAAAAACAGGTACA	1173						
Db	781	CATCCCTATGTTGGAAACAGGAGGGAGATTCATTCGGTTATCCTAAAAACAGGTACA	840						
Qy	1174	GCAATTCCTAAAGTCATCTTTAAGATGTCAGAAATAATGATGATGCTGGAAGGAAGATC	1233						
Db	841	GCAATTCCTAAAGTCATCTTTAAGATGTCAGAAATAATGATGATGCTGGAAGGAAGATC	900						
Qy	1234	ATAGATGTCATAGATAAGAACTAATCAACCTTTTGAGATTTCTATTTGAAGAGGTGAA	1293						
Db	901	ATAGATGTCATAGATAAGAACTAATCAACCTTTTGAGATTTCTATTTGAAGAGGTGAA	960						
Qy	1294	TATATTCAGAGCTGGATGGACCTCTGAGGGAATA- - -TGCTTGGTCCATCTACT	1349						
Db	961	TATATTCAGAGCTGGATGGACCTCTGAGGGAATAATGCTGCTTGGTCCATCTACT	1020						
Qy	1350	AGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCACCTGAAATTTATTCACAGT	1409						

Db	1021	AGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATCCAGT	1080						
Qy	1410	AGAAGATGATGTTATGGAAGGCAGAGACTCATTTAGTGCAGTGCCTGATTTCTGTGAGCC	1469						
Db	1081	AGAAGATGATGTTATGGAAGGCAGAGACTCATTTAGTGCAGTGCCTGATTTCTGTGAGCC	1140						
Qy	1470	ACTAAATTAATCTATGAAGAAACAAACAGACATCTGATTAATATCCATGACATCTTTCAATG	1529						
Db	1141	ACTAAATTAATCTATGAAGAAACAAACAGACATCTGATTAATATCCATGACATCTTTCAATG	1200						
Qy	1530	TTTTCCCAAAAGTCACGAAGAGGAATTTAGTGTATTTTTCCTCTGATTCGAAACAGG	1589						
Db	1201	TTTTCCCAAAAGTCACGAAGAGGAATTTAGTGTATTTTTCCTCTGATTCGAAACAGG	1260						
Qy	1590	TTTTCCGTCATTTATACAAATTAACATCTATTTTAAAGAAAGCAATATATAACGATCCAG	1649						
Db	1261	TTTTCCGTCATTTATACAAATTAACATCTATTTTAAAGAAAGCAATATATAACGATCCAG	1320						
Qy	1650	TGGTGGGCTGCTGCTCCAGTGAATTTCAAGTGCCTATCAAGAGAGAGATAGCAATTAC	1709						
Db	1321	TGGTGGGCTGCTGCTCCAGTGAATTTCAAGTGCCTATCAAGAGAGAGATAGCAATTAC	1380						
Qy	1710	CAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCAAGTTGATGAAGTCAG	1769						
Db	1381	CAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCAAGTTGATGAAGTCAG	1440						
Qy	1770	AAGGCTGATATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGT	1829						
Db	1441	AAGGCTGATATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGT	1500						
Qy	1830	CAGTTACGTAATCTCGAGAGGTGACAGGCTGACCTGACCGTGGTACTCACTTTCTTG	1889						
Db	1501	CAGTTACGTAATCTCGAGAGGTGACAGGCTGACCTGACCGTGGTACTCACTTTCTTG	1560						
Qy	1890	CTGCATCAGTCAGCACCTGTGACTTCTTTAAGTAAGTATAGTAACAGAGAAATCCACA	1949						
Db	1561	CTGCATCAGTCAGCACCTGTGACTTCTTTAAGTAAGTATAGTAACAGAGAAATCCACA	1620						
Qy	1950	CTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAGTACCCAACTTGCMAAACAAAGGA	2009						
Db	1621	CTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAGTACCCAACTTGCMAAACAAAGGA	1680						
Qy	2010	ATTTTGGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTCTGACTATATCTCTCCAGAAAT	2069						
Db	1681	ATTTTGGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTCTGACTATATCTCTCCAGAAAT	1740						
Qy	2070	TTTCTCTTTTGAAGTACTACTGGATTTTACATTTGTATGGATGCTCTACAGGCTCATGA	2129						
Db	1741	TTTCTCTTTTGAAGTACTACTGGATTTTACATTTGTATGGATGCTCTACAGGCTCATGA	1800						
Qy	2130	TCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTCAATATGTTGGTCTCTCAGGTGCA	2189						
Db	1801	TCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTCAATATGTTGGTCTCTCAGGTGCA	1860						
Qy	2190	GTTGGTGAATAATCCGGTTTAAAGGAGTCAGTATTTCCGCTTGATACCTTAGGCTCTCT	2249						
Db	1861	GTTGGTGAATAATCCGGTTTAAAGGAGTCAGTATTTCCGCTTGATACCTTAGGCTCTCT	1920						
Qy	2250	AGGTTATGTTGTTAGTATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGA	2309						
Db	1921	AGGTTATGTTGTTAGTATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGA	1980						
Qy	2310	AGGCGCTTTTAAATTAATAAATGGGTCAATAGAAATTTGAACGATCAGGTGGAAGACTCCA	2369						
Db	1981	AGGCGCTTTTAAATTAATAAATGGGTCAATAGAAATTTGAACGATCAGGTGGAAGACTCCA	2040						
Qy	2370	ATATCTAGCTTCTCCATATGATTTTCAATTTGACTTATAGATCGTGTGGGCAATCCACGGCTGTC	2429						
Db	2041	ATATCTAGCTTCTCCATATGATTTTCAATTTGACTTATAGATCGTGTGGGCAATCCACGGCTGTC	2100						
Qy	2430	CTATGAGAGATACCTCTCCCTGATGGCAATTAATCCAGAGGTCTAGATATCTTTCAGGGTTC	2489						

Db 2101 CTATGGAGGATACCTCTCCCTGATGGCAATTAATGACAGAGTTCAGATATCTTCAGGGTTGC 2160
Qy 2490 TATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATATACCGAATCGTTA 2549
Db 2161 TATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATATACCGAATCGTTA 2220
Qy 2550 TATGGGTCAACCTGACAGCAATGAACAGGCTATTAATCTAGATCTGTGGCCATCAAGC 2609
Db 2221 TATGGGTCAACCTGACAGCAATGAACAGGCTATTAATCTAGATCTGTGGCCATCAAGC 2280
Qy 2610 AGAAAAGTTCCCTCTGAACCAATCGTTTACTGTCTCTTACATGTTTCTGTGGATGAGAA 2669
Db 2281 AGAAAAGTTCCCTCTGAACCAATCGTTTACTGTCTCTTACATGTTTCTGTGGATGAGAA 2340
Qy 2670 TGTCCATTTGACATACCACTAGTATTAATCTAGATCTTGTAGGGCTTGAAGGCCATA 2729
Db 2341 TGTCCATTTGACATACCACTAGTATTAATCTAGATCTTGTAGGGCTTGAAGGCCATA 2400
Qy 2730 TGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTGAAATCGGGAGAAC 2789
Db 2401 TGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTGAAATCGGGAGAAC 2460
Qy 2790 TTATGAATCTGATCTTTTGTGACTCTGTAGAACTCTCTGTGTATACACTGGCTATTTAAACCA 2909
Db 2461 TTATGAATCTGATCTTTTGTGACTCTGTAGAACTCTCTGTGTATACACTGGCTATTTAAACCA 2520
Qy 2850 AAAAGTGTATATAATTTTGTGACTCTGTAGAACTCTCTGTGTATACACTGGCTATTTAAACCA 2909
Db 2521 AAAAGTGTATATAATTTTGTGACTCTGTAGAACTCTCTGTGTATACACTGGCTATTTAAACCA 2580
Qy 2910 ATGAGGAGTTTAAATCAACGAAACACAGAAATGATATCATCATATTTGATACCTGCCAT 2969
Db 2581 ATGAGGAGTTTAAATCAACGAAACACAGAAATGATATCATCATATTTGATACCTGCCAT 2640
Qy 2970 GTACATCTACTCTCTGAAATTAATGTGTGTCATGCGGGTCTACGGTTTGTGTAGT 3029
Db 2641 GTACATCTACTCTCTGAAATTAATGTGTGTCATGCGGGTCTACGGTTTGTGTAGT 2700
Qy 3030 AATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATTTCTGAGAGACCA 3089
Db 2701 AATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATTTCTGAGAGACCA 2760
Qy 3090 GCAATACCATGAAGATTACTAAAAA 3120
Db 2761 GCAATACCATGAAGATTACTAAAAA 2791

RESULT 14

AR631281 LOCUS AR631281 4523 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 8 from patent US 6844180.

AR631281 ACCESSION AR631281

AR631281.1 GI:59770924

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 4523)

AUTHORS Oi.S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.

TITLE Serine protease genes related to DPPIV

JOURNAL Patent: US 6844180-A 8 18-JAN-2005;

Ferring BV,;

NLX;

FEATURES Location/Qualifiers

source 1..4523

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 68.2%; Score 2128; DB 2; Length 4523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAGTGTAAAGCCTCCGAGGCCAAGGCGCTGTCTACTGCGCGCGCTGCTTCTTAGTGCCG 60
Db 1 AAGTGTAAAGCCTCCGAGGCCAAGGCGCTGTCTACTGCGCGCGCTGCTTCTTAGTGCCG 60
Qy 61 CQFTCCGCGCTGGGTTTCTCAACGCGCGCGCGCGAGGAGGACATCTGCAACACAGGACCG 120
Db 61 CQFTCCGCGCTGGGTTTCTCAACGCGCGCGCGCGAGGAGGACATCTGCAACACAGGACCG 120
Qy 121 GAGTGGAGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCCACGTCGGGACCG 180
Db 121 GAGTGGAGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCCACGTCGGGACCG 180
Qy 181 TCCGCGCGCGCGCGCGGAGGAGAAATGCAACATGCGCAGCAATGGAACACAGAACAG 240
Db 181 TCCGCGCGCGCGCGCGGAGGAGAAATGCAACATGCGCAGCAATGGAACACAGAACAG 240
Qy 241 CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Db 241 CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Qy 301 CCTAAATTGGAGCCCTTTTATGTTGAGCGGTATTTCTTGGAGTCAGCTTAAAAAGCTGCTT 360
Db 301 CCTAAATTGGAGCCCTTTTATGTTGAGCGGTATTTCTTGGAGTCAGCTTAAAAAGCTGCTT 360
Qy 361 GCCGATACAGAAATATCATGGCTCATGATGCTTAAGGCCACCAATGATTTTCATGTTT 420
Db 361 GCCGATACAGAAATATCATGGCTCATGATGCTTAAGGCCACCAATGATTTTCATGTTT 420
Qy 421 GTGAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCT 480
Db 421 GTGAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCT 480
Qy 481 GGTGAGAACAGAGAAATATCACTGTTTATTTCTGAAATTTCCAAAACTATCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAATATCACTGTTTATTTCTGAAATTTCCAAAACTATCAATAGAGCA 540
Qy 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGATCTTTTCAAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGATCTTTTCAAGGCAACACTGGACTAT 600
Qy 601 GGAATGTAATTCGAGAGAGAGAACTAATAAGAGAGAGAGAAACGCAATTTGGAACAGTCGA 660
Db 601 GGAATGTAATTCGAGAGAGAGAACTAATAAGAGAGAGAGAAACGCAATTTGGAACAGTCGA 660
Qy 661 ATTGCTTCTTACGATATCAACAGAGAGTGGACATTTCTGTTTCAAGCGGCTAGTGA 720
Db 661 ATTGCTTCTTACGATATCAACAGAGAGTGGACATTTCTGTTTCAAGCGGCTAGTGA 720
Qy 721 ATTTATCAGTAAAGATGGAGGCGCACAGGATTTACGCAACCAACCTTTAAAGGCCCAAT 780
Db 721 ATTTATCAGTAAAGATGGAGGCGCACAGGATTTACGCAACCAACCTTTAAAGGCCCAAT 780
Qy 781 CTAGTGGAAATCTAGTGTGCCCAATACGAGTGGATCCAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGGAAATCTAGTGTGCCCAATACGAGTGGATCCAAATTTATGCCCGCTGATCCA 840
Qy 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTTGGATATCTAACTGTAACACAGAGAA 900
Db 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTTGGATATCTAACTGTAACACAGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAACTGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAACTGGAAGAGATGCCAGATCA 960
Qy 961 GCTGGAGTGGCTACCTTTGTTCTTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTGGCTACCTTTGTTCTTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Qy 1021 TGTCCAAAAGCTGAAACAACTCCAGTGTGGTGTAAAATTTCTTAGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAAGCTGAAACAACTCCAGTGTGGTGTAAAATTTCTTAGAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAAACAGGAGG 1140

Db	1081	AATGATGAATCTGAGGTGGAAATATTTCATGTATACATCCCTATGTTGGAAACAGGAGG	1140
Qy	1141	GCAGATTCAATCCGTTATTCCTAAACAGGTACAGCAATCTTAAGTCATCTTTAAGATG	1200
Db	1141	GCAGATTCAATCCGTTATTCCTAAACAGGTACAGCAATCTTAAGTCATCTTTAAGATG	1200
Qy	1201	TCAGAAATAATGATTGATCTGAAGGAGGATCATAGATGTCATAGATAAGGAACATAAT	1260
Db	1201	TCAGAAATAATGATTGATCTGAAGGAGGATCATAGATGTCATAGATAAGGAACATAAT	1260
Qy	1261	CAACCTTTTGAGATCTATTTGAAGGAGTTGAATATATTGCGAGAGCTGGATGCTCT	1320
Db	1261	CAACCTTTTGAGATCTATTTGAAGGAGTTGAATATATTGCGAGAGCTGGATGCTCT	1320
Qy	1321	GAGGAAATATGCTTGCTCATCTAGATGCTCCAGACTCGCCTACAGATAGTG	1380
Db	1321	GAGGAAATATGCTTGCTCATCTAGATGCTCCAGACTCGCCTACAGATAGTG	1380
Qy	1381	TTGATCTCACTGAAATTTATTTATCCCAAGTGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Db	1381	TTGATCTCACTGAAATTTATTTATCCCAAGTGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Qy	1441	ATTGAGTCAGTGCCTGATCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC	1500
Db	1441	ATTGAGTCAGTGCCTGATCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC	1500
Qy	1501	TGGATAAATATCCATGATCATCTTTTCATGTTTTCCTCCAAAGTCACGAAGAGGAAATGAG	1560
Db	1501	TGGATAAATATCCATGATCATCTTTTCATGTTTTCCTCCAAAGTCACGAAGAGGAAATGAG	1560
Qy	1561	TTTATTTTTCCTCTGAAATCAAAACAGGTTTCCTGTCATTTATATCAAAATTTACATCTTT	1620
Db	1561	TTTATTTTTCCTCTGAAATCAAAACAGGTTTCCTGTCATTTATATCAAAATTTACATCTTT	1620
Qy	1621	TTAAAGGAAAGCAAAATATAAAGATCCAGTGTGGGCTGCTGCTCCAAAGTATTTCAAG	1680
Db	1621	TTAAAGGAAAGCAAAATATAAAGATCCAGTGTGGGCTGCTGCTCCAAAGTATTTCAAG	1680
Qy	1681	TGTCCTATCAAGAGAGATAGCAATTTACAGTGTGGAATGGGAGTTCTTGGCCGGCAT	1740
Db	1681	TGTCCTATCAAGAGAGATAGCAATTTACAGTGTGGAATGGGAGTTCTTGGCCGGCAT	1740
Qy	1741	GGATCTAATATCCAAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1800
Db	1741	GGATCTAATATCCAAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1800
Qy	1801	TCCCTTTAGAGCATCACCTGTACGTAGTACGTAGTACGTAGTACGTAGTACGTAGTACGT	1860
Db	1801	TCCCTTTAGAGCATCACCTGTACGTAGTACGTAGTACGTAGTACGTAGTACGTAGTACGT	1860
Qy	1861	CTGACTGACGGTGGTACTCAATCTTGTGCTGATCAGTACGTAGTACGTAGTACGTAGTAC	1920
Db	1861	CTGACTGACGGTGGTACTCAATCTTGTGCTGATCAGTACGTAGTACGTAGTACGTAGTAC	1920
Qy	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG	1980
Db	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG	1980
Qy	1981	GAGATGACCCCACTTGCACAAACAAAGGAAATTTTGGGCGCCACCAATTTTGGATTCAGGCT	2040
Db	1981	GAGATGACCCCACTTGCACAAACAAAGGAAATTTTGGGCGCCACCAATTTTGGATTCAGGCT	2040
Qy	2041	CCTCTTCTGACTATACCTCCAGAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT	2100
Db	2041	CCTCTTCTGACTATACCTCCAGAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT	2100
Qy	2101	TTGTATGGAGTCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTG	2160
Db	2101	TTGTATGGAGTCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTG	2160
Qy	2161	CTGTTTCATATATGGTGTGTC	2179
Db	2161	CTGTTTCATATATGGTGTGTC	2179
RESULT 15			
AX608731			
LOCUS			
DEFINITION	Sequence 8 from Patent WO0231134.	4523 bp	DNA linear PAT 17-FEB-2003
ACCESSION	AX608731		
VERSION	AX608731.1	GI:28404301	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.		
TITLE	Novel serine protease genes related to dppiv		
JOURNAL	Patent: WO 0231134-A 8 18-APR-2002;		
	Ferring BV (NL)		
FEATURES	Location/Qualifiers		
source	1..4523		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	68.2%;	Score 2128;	DB 2; Length 4523;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 2178;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	AAGTGCTAAAGCCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG	60
Db	1	AAGTGCTAAAGCCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG	60
Qy	61	CGTTCCGCGCTGGGTTGTCAACGCGCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120
Db	61	CGTTCCGCGCTGGGTTGTCAACGCGCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120
Qy	121	GAGTGGAGGCG	180
Db	121	GAGTGGAGGCG	180
Qy	181	TCCG	240
Db	181	TCCG	240
Qy	241	CTGGTGCTTCAGATATTTGAAACTCGGACTCTGAGGAGATATTGAATCAGAGATCGG	300
Db	241	CTGGTGCTTCAGATATTTGAAACTCGGACTCTGAGGAGATATTGAATCAGAGATCGG	300
Qy	301	CCTAAATTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAGAGCTGCTT	360
Db	301	CCTAAATTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAGAGCTGCTT	360
Qy	361	GCGGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTT	420
Db	361	GCGGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTT	420
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTATTACCTTTGCCATGCT	480
Db	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTATTACCTTTGCCATGCT	480
Qy	481	GGTGAGAAACAGAGAAATACACTGTTTATTTCTGAAATTTCCAAAACTATCAATAGAGCA	540
Db	481	GGTGAGAAACAGAGAAATACACTGTTTATTTCTGAAATTTCCAAAACTATCAATAGAGCA	540
Qy	541	GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Db	541	GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Qy	601	GGAAATGATTTCTCGAGAGAAAGAACTATTAGAGAAAGAAACCGCATTTGGAACAGTCGGA	660
Db	601	GGAAATGATTTCTCGAGAGAAAGAACTATTAGAGAAAGAAACCGCATTTGGAACAGTCGGA	660

Db 601 GGAATGATATTCTCGAAGAGAACTATTAGAGAAAGAAAACGCAATTGGAAACAGTCGGA 660
Qy 661 ATTGCTCTTACGATTATCACCAAGAGAGTGAACATTTCTGTGTTCAAGCCGGTAGTGA 720
Db 661 ATTGCTCTTACGATTATCACCAAGAGAGTGAACATTTCTGTGTTCAAGCCGGTAGTGA 720
Qy 721 ATTATACGTAAGAGATGGAGGCCACAAGGATTTACGCAACAACTTTAAGGCCCAAT 780
Db 721 ATTATACGTAAGAGATGGAGGCCACAAGGATTTACGCAACAACTTTAAGGCCCAAT 780
Qy 781 CTAGTGGAACTAGTTGTCCTCCCAACATACGAGTGCATCCAAATTTATGCCCCGCTGATCCA 840
Db 781 CTAGTGGAACTAGTTGTCCTCCCAACATACGAGTGCATCCAAATTTATGCCCTGCTGATCCA 840
Qy 841 GACTGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACACAGAGAA 900
Db 841 GACTGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACACAGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCA 960
Qy 961 GCTGGAGTGCCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020
Db 961 GCTGGAGTGCCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020
Qy 1021 TGTCCTAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTAGAAATTTCTATATGAAGAA 1080
Db 1021 TGTCCTAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTAGAAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCTATGTTGGAAACAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCTATGTTGGAAACAGGAGG 1140
Qy 1141 GCAGATTCATTCGGTTATCTCTAAACAGGTAACAGGATCTTAAAGTCACTTTTAAAGATG 1200
Db 1141 GCAGATTCATTCGGTTATCTCTAAACAGGTAACAGGATCTTAAAGTCACTTTTAAAGATG 1200
Qy 1201 TCAGAAATATGATGATGCTGAAGGAGGATCATAGATGTCATAGATAAGAACTAAT 1260
Db 1201 TCAGAAATATGATGATGCTGAAGGAGGATCATAGATGTCATAGATAAGAACTAAT 1260
Qy 1261 CAACCTTTTGAGATTTCTATTGTAAGGAGTTGAATATTTGCCAGAGCTGGATGACTCCT 1320
Db 1261 CAACCTTTTGAGATTTCTATTGTAAGGAGTTGAATATTTGCCAGAGCTGGATGACTCCT 1320
Qy 1321 GAGGGAATATGCTTGGTCCATCTCTAGATGCTGCCAGCTCGCTACAGATAGTG 1380
Db 1321 GAGGGAATATGCTTGGTCCATCTCTAGATGCTGCCAGACTCGCTACAGATAGTG 1380
Qy 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAGATGATGTTATGGAAGGAGAGACTC 1440
Db 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAGATGATGTTATGGAAGGAGAGACTC 1440
Qy 1441 ATTGAGTCAGGCTGATTTCTGAGCCCACTAATTTATCTATGAAGAAACACAGACATC 1500
Db 1441 ATTGAGTCAGGCTGATTTCTGAGCCCACTAATTTATCTATGAAGAAACACAGACATC 1500
Qy 1501 TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTCAAGAGAGAAATTTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTCAAGAGAGAAATTTGAG 1560
Qy 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCGGTCATTTATACAAATATCACTATT 1620
Db 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCGGTCATTTATACAAATATCACTATT 1620
Qy 1621 TTAAGGAAGCAATATAAAGATCCAGTGGTGGCTGCTCTCCAAAGTATTTCAAG 1680
Db 1621 TTAAGGAAGCAATATAAAGATCCAGTGGTGGCTGCTCTCCAAAGTATTTCAAG 1680
Qy 1681 TGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Db 1681 TGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740

Qy 1741 GGATCTAATATCCAAAGTTGATGAAGCTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Db 1741 GGATCTAATATCCAAAGTTGATGAAGCTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Qy 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCGAAGAGGTGACAAGG 1860
Db 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCGAAGAGGTGACAAGG 1860
Qy 1861 CTGACTGACCGTGGCTACTCAGATTTCTGCTGTCATCAGTCAGCACTGTGATCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCAGATTTCTGCTGTCATCAGTCAGCACTGTGATCTTTATA 1920
Qy 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCTTACCAAGCTATCAAGTCT 1980
Db 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCTTACCAAGCTATCAAGTCT 1980
Qy 1981 GAGATGACCAACTTTGCAAAACAAAGAAATTTGGGCCACCATTTGGATTACAGAGT 2040
Db 1981 GAGATGACCAACTTTGCAAAACAAAGAAATTTGGGCCACCATTTGGATTACAGAGT 2040
Qy 2041 CCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACA 2100
Db 2041 CCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACA 2100
Qy 2101 TTGATGGGATGCTCTACAGGCTCATGATCTACAGGCTGGAAGAAATATCTCTACTGTG 2160
Db 2101 TTGATGGGATGCTCTACAGGCTCATGATCTACAGGCTGGAAGAAATATCTCTACTGTG 2160
Qy 2161 CTGTTCTATATGTTGTC 2179
Db 2161 CTGTTCTATATGTTGTC 2179

RESULT 16
AR631287
LOCUS AR631287 4676 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 20 from patent US 6844180.
ACCESSION AR631287
VERSION AR631287.1 GI:59770934
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 4676)
AUTHORS OI, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 20 18-JAN-2005;
Ferring BV;;
NLX;

FEATURES
source
1. 4676
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 68.2%; Score 2128; DB 2; Length 4676;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AGTGTAAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCTGCTCTTTAGTGGCG 60
Db 1 AGTGTAAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCTGCTCTTTAGTGGCG 60
Qy 61 CGTTGCGCGCTGGGTGTGTACCGGCGCGCGCGCGAGGAGCCACTGCAACACAGGACCG 120
Db 61 CGTTGCGCGCTGGGTGTGTACCGGCGCGCGCGCGAGGAGCCACTGCAACACAGGACCG 120
Qy 121 GAGTGGAGCGCGCAGCATGAAAGCGCGCGCGCGCTCCATAGCCAGTGGGACCG 180
Db 121 GAGTGGAGCGCGCAGCATGAAAGCGCGCGCGCTCCATAGCCAGTGGGACCG 180
Qy 181 TCCGGCGCGCGCGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACAGAACAG 240

Db 181 ||||| TCCGGCGGGCGGGGGAAGGAAAATGCAACATGSCACAGCAATGGAACAGACAG 240
Qy 241 ||||| CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTCGAATCAGAGATCGG 300
Db 241 ||||| CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTCGAATCAGAGATCGG 300
Qy 301 ||||| CTTAAATTTGGAGCCTTTTATGTTGAGGGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360
Db 301 ||||| CTTAAATTTGGAGCCTTTTATGTTGAGGGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360
Qy 361 ||||| GCCGATACAGAAATATCATGCTCATCATGATGGCTTAAGCCACCAATGATTTTCATGTTT 420
Db 361 ||||| GCCGATACAGAAATATCATGCTCATCATGATGGCTTAAGCCACCAATGATTTTCATGTTT 420
Qy 421 ||||| GTGAGAGGAATGATCCGATGGAACCTCATTCAGACAGAAATCTATPACCTTGCCATGTCT 480
Db 421 ||||| GTGAGAGGAATGATCCGATGGAACCTCATTCAGACAGAAATCTATPACCTTGCCATGTCT 480
Qy 481 ||||| GGTGAGAACAGAGAAATATACATGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Db 481 ||||| GGTGAGAACAGAGAAATATACATGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Qy 541 ||||| GCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 ||||| GCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 ||||| GGAATGTAATTCAGAGAGAGAACTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 ||||| GGAATGTAATTCAGAGAGAGAGAACTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 661 ||||| ATTGCTCTTACGATATACCAAGGAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGGA 720
Db 661 ||||| ATTGCTCTTACGATATACCAAGGAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGGA 720
Qy 721 ||||| ATTTATCACGTAAGAGATGAGGGCCACAAGGATTTAGCAACACCTTTTAAGGCCCAAT 780
Db 721 ||||| ATTTATCACGTAAGAGATGAGGGCCACAAGGATTTAGCAACACCTTTTAAGGCCCAAT 780
Qy 781 ||||| CTAGTGAAACTAGTTGTCCCAACATACCGATGGATCCAAATTTATGCCCCTGATCCA 840
Db 781 ||||| CTAGTGAAACTAGTTGTCCCAACATACCGATGGATCCAAATTTATGCCCCTGATCCA 840
Qy 841 ||||| GACTGGATTCCTTTATACATAGCAACGATATTTGGATCTTAACATCGTAAACAGAGAA 900
Db 841 ||||| GACTGGATTCCTTTATACATAGCAACGATATTTGGATCTTAACATCGTAAACAGAGAA 900
Qy 901 ||||| GAAAGGAGACTCACTTATGTCACATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960
Db 901 ||||| GAAAGGAGACTCACTTATGTCACATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960
Qy 961 ||||| GCTGGAGTGCCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTCCTGGCTATTTGGTGG 1020
Db 961 ||||| GCTGGAGTGCCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTCCTGGCTATTTGGTGG 1020
Qy 1021 ||||| TGTCCAAAAGCTGAAACAACTCCCAAGTGGTGAATTTCTAGAAATTTCTATATGAGAA 1080
Db 1021 ||||| TGTCCAAAAGCTGAAACAACTCCCAAGTGGTGAATTTCTAGAAATTTCTATATGAGAA 1080
Qy 1081 ||||| AATGATGAATCTGAGGTGAAATTTATGTTTACATCCCTATGTTGGAAACAAAGGAGG 1140
Db 1081 ||||| AATGATGAATCTGAGGTGAAATTTATGTTTACATCCCTATGTTGGAAACAAAGGAGG 1140
Qy 1141 ||||| GCAGATTCATTCCTGTTATCCTAAAAACAGTACAGCAATCCTAAAGTCACTTTTAAGATG 1200
Db 1141 ||||| GCAGATTCATTCCTGTTATCCTAAAAACAGTACAGCAATCCTAAAGTCACTTTTAAGATG 1200
Qy 1201 ||||| TCAGAAATTAATGATGCTGAAGCAAGGATCATAGATGTCATAGATAGGAACTAATTT 1260
Db 1201 ||||| TCAGAAATTAATGATGCTGAAGCAAGGATCATAGATGTCATAGATAGGAACTAATTT 1260
Qy 1261 ||||| CAACTTTTGGAGATCTATTTGAAGGAGTTCGAATATATTTGCCAGAGCTGGATGCTCT 1320
Db 1261 ||||| CAACTTTTGGAGATCTATTTGAAGGAGTTCGAATATATTTGCCAGAGCTGGATGCTCT 1320

Db 1261 ||||| CAACTTTTGGAGATCTATTTGAAGGAGTTCGAATATATTTGCCAGAGCTGGATGCTCT 1320
Qy 1321 ||||| GAGGAAAAATATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTG 1380
Db 1321 ||||| GAGGAAAAATATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTG 1380
Qy 1381 ||||| TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATTTGGAAGAGGAGAGACTC 1440
Db 1381 ||||| TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATTTGGAAGAGGAGAGACTC 1440
Qy 1441 ||||| ATTGAGTCAGTCCCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAAAACAGACATC 1500
Db 1441 ||||| ATTGAGTCAGTCCCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAAAACAGACATC 1500
Qy 1501 ||||| TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAAAGTCACGAAGAGGAAATTTGAG 1560
Db 1501 ||||| TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAAAGTCACGAAGAGGAAATTTGAG 1560
Qy 1561 ||||| TTTATTTTTTGGCTCTGATGCAAAACAGGTTTCGGTCAATTTATACAAATTTACATCTATT 1620
Db 1561 ||||| TTTATTTTTTGGCTCTGATGCAAAACAGGTTTCGGTCAATTTATACAAATTTACATCTATT 1620
Qy 1621 ||||| TTAAGAGAAAGCAATATTAACGATCCAGTGGTGGGCTGCTCTCCAGTGATTTCAAG 1680
Db 1621 ||||| TTAAGAGAAAGCAATATTAACGATCCAGTGGTGGGCTGCTCTCCAGTGATTTCAAG 1680
Qy 1681 ||||| TGTCTCTATCAAGAGAGAGATAGCAATTTACAGTGGTGGATGGGAAGTTCTTGGCCGGCAT 1740
Db 1681 ||||| TGTCTCTATCAAGAGAGAGATAGCAATTTACAGTGGTGGATGGGAAGTTCTTGGCCGGCAT 1740
Qy 1741 ||||| GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCAACAAAGAC 1800
Db 1741 ||||| GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCAACAAAGAC 1800
Qy 1801 ||||| TCCCTCTTGAAGCATCACTGTAGTAGTCAGTTACGTAATCTCGAGAGGTGCAAGG 1860
Db 1801 ||||| TCCCTCTTGAAGCATCACTGTAGTAGTCAGTTACGTAATCTCGAGAGGTGCAAGG 1860
Qy 1861 ||||| CTGACTGACCGTGGCTACTCACATTTCTGCTGATCAGTCAGCACTGTGACTTCTTTATA 1920
Db 1861 ||||| CTGACTGACCGTGGCTACTCACATTTCTGCTGATCAGTCAGCACTGTGACTTCTTTATA 1920
Qy 1921 ||||| AGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCTCT 1980
Db 1921 ||||| AGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTCTCTCTTTTACAAGCTATCAAGTCTCT 1980
Qy 1981 ||||| GAAGATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTCACAGGT 2040
Db 1981 ||||| GAAGATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTCACAGGT 2040
Qy 2041 ||||| CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA 2100
Db 2041 ||||| CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA 2100
Qy 2101 ||||| TTGTATGGGATGCTTACAGCTCATGATCTACAGCTGGAAAGAAATATCTTACTGTG 2160
Db 2101 ||||| TTGTATGGGATGCTTACAGCTCATGATCTACAGCTGGAAAGAAATATCTTACTGTG 2160
Qy 2161 ||||| CTGTTTCATATATGGTGTCT 2179
Db 2161 ||||| CTGTTTCATATATGGTGTCT 2179

RESULT: 17

AX608743

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Sequence 20 from Patent WO0231134.

AX608743.1 GI:28404307

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AX608743

4676 bp

DNA

linear

PAT 17-FEB-2003

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE
1 Q1.S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
AUTHORS
TITLE Novel serine protease genes related to dplv
JOURNAL Patent: WO 0231134-A 20 18-APR-2002;
Ferring BV (NL)

FEATURES
source Location/Qualifiers
1..4676
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 68.2%; Score 2128; DB 2; Length 4676;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAGTCTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCGCGCGTCTTCTTAGTGGCG 60
Db 1 AAGTCTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCGCGCGTCTTCTTAGTGGCG 60
QY 61 CGTTCGCCCGCTGGTGTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CGTTCGCCCGCTGGTGTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 GAGTGGAGCG 180
Db 121 GAGTGGAGCG 180
QY 181 TCCGGCG 240
Db 181 TCCGGCG 240
QY 241 CTGGGGTGTGAGATATTTGAACTCGGACTGTGAGGAGATATTTGAACTCAAGGATCGG 300
Db 241 CTGGGGTGTGAGATATTTGAACTCGGACTGTGAGGAGATATTTGAACTCAAGGATCGG 300
QY 301 CCTAAATGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGCTTAAAGCTGCTT 360
Db 301 CCTAAATGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGCTTAAAGCTGCTT 360
QY 361 GCCGATACAGAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTCATGTT 420
Db 361 GCCGATACAGAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTCATGTT 420
QY 421 GTGAGAGGAATCATCCAGATGACCTCATTCAGACAGAACTTATACCTTGGCCATGTCT 480
Db 421 GTGAGAGGAATCATCCAGATGACCTCATTCAGACAGAACTTATACCTTGGCCATGTCT 480
QY 481 GGTGAGAACAGAGAAATACATGTTTATTCGAAATTCGAAACTATCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAATACATGTTTATTCGAAATTCGAAACTATCAATAGAGCA 540
QY 541 GCAGTCTTAATGCTCTTGGAGCCTCTTTTGGATCTTTTTCAGGCAACACTTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTTGGAGCCTCTTTTGGATCTTTTTCAGGCAACACTTGGACTAT 600
QY 601 GGAATGTATTCTCCAGAGAGAGAACTATTAAGAGAGAGAAACCGATTTGGAACAGTCGGA 660
Db 601 GGAATGTATTCTCCAGAGAGAGAACTATTAAGAGAGAGAAACCGATTTGGAACAGTCGGA 660
QY 661 ATTGCTTCTTACGATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 ATTGCTTCTTACGATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 ATTTATACAGTAAAG 780
Db 721 ATTTATACAGTAAAG 780
QY 781 CTAGTGGAAACTAGTGTGTCACAAATATCGAGATGATCCAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGGAAACTAGTGTGTCACAAATATCGAGATGATCCAAATTTATGCCCGCTGATCCA 840

QY 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAACCAAGAGAA 900
Db 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAACCAAGAGAA 900
QY 901 GAAAGGAGACTCACTTATGTGCAAAATAGCTAGCCAAACATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGCAAAATAGCTAGCCAAACATGGAAGAGATGCCAGATCA 960
QY 961 GCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
QY 1021 TGTCCAAAGAGCTGAAACCAACTCCCAAGTGGTGAATAATCTTAGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAGAGCTGAAACCAACTCCCAAGTGGTGAATAATCTTAGAATTTCTATATGAAGAA 1080
QY 1081 AATGATGAATCTGAGGTGGAAATTTATCATGTTTACATCCCTCTATGTTGGAACCAAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGGAAATTTATCATGTTTACATCCCTCTATGTTGGAACCAAGGAGG 1140
QY 1141 GCAGATTCATTCCGTTTATCTTAAACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAGATG 1200
Db 1141 GCAGATTCATTCCGTTTATCTTAAACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAGATG 1200
QY 1201 TCAGAAATATGATGATGCTGAAGAGAGATCATAGATGTCATAGATGATAGATAAGGAACTAATT 1260
Db 1201 TCAGAAATATGATGATGCTGAAGAGAGATCATAGATGTCATAGATGATAGATAAGGAACTAATT 1260
QY 1261 CAACCTTTTGAATTTCTTAAAGAGAGATGTAATATATTTGCGAGAGCTGGATGACTCCT 1320
Db 1261 CAACCTTTTGAATTTCTTAAAGAGAGATGTAATATATTTGCGAGAGCTGGATGACTCCT 1320
QY 1321 GAGGAAATATGCTTTGGTCCATCTCTACTAGATGCTCCAGACTCGCTTACAGATAGTG 1380
Db 1321 GAGGAAATATGCTTTGGTCCATCTCTACTAGATGCTCCAGACTCGCTTACAGATAGTG 1380
QY 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGAGAGAGAGAGAG 1440
Db 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGAGAGAGAGAGAG 1440
QY 1441 ATTGAGTCAGTGGCTGATCTGAGCGCACTAAATATCTATGAAAGAAACCAACAGACATC 1500
Db 1441 ATTGAGTCAGTGGCTGATCTGAGCGCACTAAATATCTATGAAAGAAACCAACAGACATC 1500
QY 1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTCCTCCAAAGTCAGAGAGAGAGAGAGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTCCTCCAAAGTCAGAGAGAGAGAGAGAG 1560
QY 1561 TTTATTTTGGCTCTGAAATGCAAAACAGAGTTTCCCGTCATTATATACAAAATTACATCTATT 1620
Db 1561 TTTATTTTGGCTCTGAAATGCAAAACAGAGTTTCCCGTCATTATATACAAAATTACATCTATT 1620
QY 1621 TTAAGAGAAAGCAAAATATAAAGCATCAAGTGGTGGCTGCTCTCCAAAGTATTCAAG 1680
Db 1621 TTAAGAGAAAGCAAAATATAAAGCATCAAGTGGTGGCTGCTCTCCAAAGTATTCAAG 1680
QY 1681 TGTCTCTATCAAGAGAGAGATAGCAATTTACAGTGGTGAATGGAAGTTCTTGGCCGGCAT 1740
Db 1681 TGTCTCTATCAAGAGAGAGATAGCAATTTACAGTGGTGAATGGAAGTTCTTGGCCGGCAT 1740
QY 1741 GGATCTAATATCCAAAGTTGATGAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGC 1800
Db 1741 GGATCTAATATCCAAAGTTGATGAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGC 1800
QY 1801 TCCCTTTTAGAGCATCACTGTACGTAGTCACTTAATCTGAGAGAGAGAGAGAGAGAGAG 1860
Db 1801 TCCCTTTTAGAGCATCACTGTACGTAGTCACTTAATCTGAGAGAGAGAGAGAGAGAGAG 1860
QY 1861 CTGACTGACCGTGGCTACTCACTTTCTGTCAGTCACTGAGTCACTGAGTCTTCTTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACTTTCTGTCAGTCACTGAGTCACTGAGTCTTCTTTTATA 1920

```
QY 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCT 1980
DB 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCT 1980
QY 1981 GAAAGATGACCCCAACTTGCACCAAAAGGAATTTTGGCCACACATTTTGAATTCAGAGGT 2040
DB 1981 GAAAGATGACCCCAACTTGCACCAAAAGGAATTTTGGCCACACATTTTGAATTCAGAGGT 2040
QY 2041 CCTCTTCTGACTACTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACA 2100
DB 2041 CCTCTTCTGACTACTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACA 2100
QY 2101 TTGTATGGAGCTCTACAGGCTCATGATCTACAGCTTGGAAAGAAATATCTACTGTG 2160
DB 2101 TTGTATGGAGCTCTACAGGCTCATGATCTACAGCTTGGAAAGAAATATCTACTGTG 2160
QY 2161 CTGTTTCATATATGTTGTC 2179
DB 2161 CTGTTTCATATATGTTGTC 2179

RESULT 18
AX405770 2842 bp DNA linear PAT 14-JUN-2002
SEQUENCE 185 from Patent WO222660.
AX405770
AX405770.1 GI:21438979
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1
Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, P.,
Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.
Novel nucleic acids and polypeptides
Patent: WO 0222660-A 185 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
Source Location/Qualifiers
1..2842
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
234..2582
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD34877.1"
/db_xref="GI:21438980"
/translation="
LKKLLADPKFKHGMKAKPHDFMVKRNDPDPGPHSDRIYLLAMSGENRENTLFYSEI
PTFLNRAALWLSWPKLLDLFQATLDYGMYSREELLREKRIGTVGASDYDHQSG
TEFLQAGSGITHVKGQPGFTQQLRPNLVETSCPNRMDPKLPADPDWIAPIHNS
DIWISNIVTREERLTYHNELANMEEDARSAGVATFVLQEBFDYRYGVMWCPKAETI
PSGKILRIEENDESEVEIHTVSPMLERRADSFYKPTGTANPKVTFKMSIMI
DARGKLIIDVDEKLIQPEILFEGVEYIARAGWTPGKYAWSILLDRSQRLQIWLIS
FELFIPEDVDMERORLIESVPDSVTLPIIYEETTDIWINIHDIFHVPQSEHEIEF
IPASECKTGFPHLYKITSILKESKYRSGGLPAPSPDFKCPIKEBIATISGHEWELGR
HGSNIQVDEVRRLVYFPGTKDSPLEHLLVYVSYNPGVEVRLTDRGYSHSCISOHCD
FISKYNQRNPKHCVSLYKLSPEDDPTCKTFEAWATILDSAGPLPDVTPPEIFSFES
TGTFTLGMLYKPHDLQPKKYPVLFIYGGPQVQIAGAPVTLMIYDGTGYERYMGH
PDQNEQGYLLGSVAMQAEKFPSPBNRLLLHGLFDENVHFAHTSILLSFLVRAGKPYD
LQIYVQSRHSIRVPESGHEVELHLLYLOENLGRISLAALKVI"

Query Match 67.8%; Score 2116; DB 2; Length 2842;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2166; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 GCCAAGCGCGTGTACTGCGCGCTGCTTCTTAGTGGCGGTTCGCGCGCTGGGTGT 79
DB 40 GCCAAGCGCGTGTACTGCGCGCTGCTTCTTAGTGGCGGTTCGCGCGCTGGGTGT 99
```

Db	1180	CTAAACAGGTACAGCAATCCTTAAGTCACTTTTAAGATGTCAGAAATAATGATTGATG	1239
Qy	1220	CTGAAGGAGGATCATAGATGTCATAGATAGAGAACTTAATCAACCTTTTGAGATTCCTAT	1279
Db	1240	CTGAAGGAGGATCATAGATGTCATAGATAGAGAACTTAATCAACCTTTTGAGATTCCTAT	1299
Qy	1280	TTGAAGGAGTGAATATATATGTCAGAGCTGATGGAATCTCTGAGGAAATAATGCTTGGT	1339
Db	1300	TTGAAGGAGTGAATATATATGTCAGAGCTGATGGAATCTCTGAGGAAATAATGCTTGGT	1359
Qy	1340	CCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAATAT	1399
Db	1360	CCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAATAT	1419
Qy	1400	TTATCCCAAGTGAAGATGATGTTATGAAAAGGAGAGATCATGATGAGTGCCTGATT	1459
Db	1420	TTATCCCAAGTGAAGATGATGTTATGAAAAGGAGAGATCATGATGAGTGCCTGATT	1479
Qy	1460	GTGTGACGCGCACTAATATCTATGAGAAACACAGACATCTGATAAATATCCATGACA	1519
Db	1480	GTGTGACGCGCACTAATATCTATGAGAAACACAGACATCTGATAAATATCCATGACA	1539
Qy	1520	TCCTTCATGTTTTTCCCAAGTCACGAAGAGGAAATGAGTTTATTTTGGCTCTGAT	1579
Db	1540	TCCTTCATGTTTTTCCCAAGTCACGAAGAGGAAATGAGTTTATTTTGGCTCTGAT	1599
Qy	1580	GCAAAACAGGTTTCGCTCATTTATACAAATTAATCATTTTAAAGGAAAGCAATATA	1639
Db	1600	GCAAAACAGGTTTCGCTCATTTATACAAATTAATCATTTTAAAGGAAAGCAATATA	1659
Qy	1640	AACGATCCAGTGGTGGCTGCTGCTCAAGTGATTTCAAGTGCTCTATCAAGAGGACA	1699
Db	1660	AACGATCCAGTGGTGGCTGCTGCTCAAGTGATTTCAAGTGCTCTATCAAGAGGACA	1719
Qy	1700	TAGCAATACAGTGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAAGTTG	1759
Db	1720	TAGCAATACAGTGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAAGTTG	1779
Qy	1760	ATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGATCTCCCTTTAGAGCATCAC	1819
Db	1780	ATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGATCTCCCTTTAGAGCATCAC	1839
Qy	1820	TGTACGTAGTCAGTTACGTAATCTGTGAGAGGTGACAGGCTCACTGACCGTGGCTACT	1879
Db	1840	TGTACGTAGTCAGTTACGTAATCTGTGAGAGGTGACAGGCTCACTGACCGTGGCTACT	1899
Qy	1880	CACATTCCTGCTGATCAGTCAGCCTGCTGCTTTTAAAGTAAAGTATAGTAACGAGA	1939
Db	1900	CACATTCCTGCTGATCAGTCAGCCTGCTGCTTTTAAAGTAAAGTATAGTAACGAGA	1959
Qy	1940	AGAATCCACACTGTGTGCTTTCACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCA	1999
Db	1960	AGAATCCACACTGTGTGCTTTCACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCA	2019
Qy	2000	AAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGAGGTCCTCTTCTGACTATCTC	2059
Db	2020	AAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGAGGTCCTCTTCTGACTATCTC	2079
Qy	2060	CTCCAGAAATTTCTTTTGAAGTACTACTGGAATTTACATGATGATGGATGCTCTACA	2119
Db	2080	CTCCAGAAATTTCTTTTGAAGTACTACTGGAATTTACATGATGATGGATGCTCTACA	2139
Qy	2120	AGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTGCTGTTCATATATGGTGGTC	2179
Db	2140	AGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTGCTGTTCATATATGGTGGTC	2199
Qy	2180	CTCAGGT 2186	
Db	2200	CTCAGGT 2206	

BC030688
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL
REMARK
COMMENT
CONSTRM
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL
REMARK
COMMENT

BC030688 3125 bp mRNA linear PRI 19-JAN-2006
Homo sapiens dipeptidyl-peptidase 8, transcript variant 4, mRNA
(cDNA clone MGC:26191 IMAGE:4822550), complete cds.
BC030688
BC030688.2 GI:34190028
MGC.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 3125)
Straussberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Loquellano,N.A., Peters,G.J.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
McKernan,K.J., Hale,S.S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Kravinsky,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3125)
NIH MGC Project
Direct Submission
Submitted (24-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
On Aug 25, 2003 this sequence version replaced gi:21265132.
Contact: MGC help desk
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Wiklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahney, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 33 Row: d Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 37577090.
Location/Qualifiers
1. 3125
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:26191 IMAGE:4822550"
/tissue_type="Testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescriptR"
1. 3125
/gene="DPP8"

/note="synonyms: FLJ20283, FLJ14920, MGC26191, MSTP141, DP8, DRP1"			
/db_xref="GeneID:54878"			
/db_xref="HGNC:16490"			
/db_xref="IMGT/GENE-DB:16490"			
/db_xref="MIM:606819"			
319..2862			
/gene="DPP8"			
/codon_start=1			
/product="dipeptidyl peptidase 8, isoform 4"			
/protein_id="AAH30688.3"			
/db_xref="GI:54114905"			
/db_xref="GeneID:54878"			
/db_xref="HGNC:16490"			
/db_xref="IMGT/GENE-DB:16490"			
/db_xref="MIM:606819"			
/translation="MWKSEQMKIKSGKCNMAAAMETQLGVEIFETADCEENIESQD RPLKPFVVERYSQKLKLLADTRKYHGMKAPDFMFKRNDPDGPHSDRIYYL AMSENGRENTFYSEIPKTIINRAALVMSLWKPLDFOATLDYGMVSRBELRERYL IGTVGIASDYHQSGTFLFQAGSGIYHVKGQPGQFTQOPLRPNLVSTCPNIRMDP KLCPADPDWIAFTSHNDIWSNIVTRERRLLTYVHNELANNEEDARSAGVATVLOEE FDRYGVWPCMAETTPSGGKILRLIYBENDESEVEIIVHTSPMLRFRADSPKRYKPT GTANPKVTFKMEIMIDAGRIIDVIDKELIQPEILLPEGVETIARAGWTPEGKRYWS ILLDRSOTRLOIVLISPELFIPEVDVNERQRILIESVPDSVTPLIYEEETDIWINIH DIHFVFPQSHBEIEIFIASECKTGFRLYKITSILKESYKRSRSGGLPAPSPKCP KBIATITSGEWEVLRGHSNIOQVSRVRLVFEKTKDPSLEHLHYVSVYVNRGEVEL TRDGYSHSCCISQCHDFPFIISKYNKPHCVSLYKLSPEDDPTCKTKEMFATYILSL GLPUDYTPPEIFSTGFTLYGMLYKPHLDQPKKYPTVLFIYGGPQVLVNNRFK GVKYPRLNTLASLGYVWVLDNRGSHRGLKFEKAPYKMWALAGAPVLNLIIFIDIGY TERYMGHPDNQGGYVLGSAVMAQEKFPSEPNRLLLLGHFDENVHFAHTSILLISFLV RACKPYDLQIYQERHSIRVPESGHEHYELHLLHYLQENLGSRIAAALKVI"			
ORIGIN			
Query Match		66.7%; Score 2081; DB 5; Length 3125;	
Best Local		Similarity 99.9%; Pred. No. 0;	
Matches 213;		Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	201	AGGAAATGCAACATGCGCAGCAGCAATGGAACAGAACAGACAGCTGGGTGTGAGATATTGA	260
DB	354	AGGAAATGCAACATGCGCAGCAGCAATGGAACAGAACAGACAGCTGGGTGTGAGATATTGA	413
QY	261	AATCGCGACTGTGAGGAGAAATTTGAATCAGAGATCGGCTAAATTTGGAGCCTTTT	320
DB	414	AATCGCGACTGTGAGGAGAAATTTGAATCAGAGATCGGCTAAATTTGGAGCCTTTT	473
QY	321	TGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAAAAATCA	380
DB	474	TGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAAAAATCA	533
QY	381	TGCTACATGATGGCTAAGGACACCATGATTTTATGTTTGTGAAGAGAAATGATCCAGA	440
DB	534	TGCTACATGATGGCTAAGGACACCATGATTTTATGTTTGTGAAGAGAAATGATCCAGA	593
QY	441	TGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCTGCTGTGAGAACAGAGAAAAATAC	500
DB	594	TGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCTGCTGTGAGAACAGAGAAAAATAC	653
QY	501	ACTGTTTATTTCTGAAATTTCCAAACATCAATAGAGCAGCAGCTTTAATGCTCTCTTG	560
DB	654	ACTGTTTATTTCTGAAATTTCCAAACATCAATAGAGCAGCAGCTTTAATGCTCTCTTG	713
QY	561	GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGA	620
DB	714	GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGA	773
QY	621	AGAACTATTAAAGAGAAAGAAACGCAATGGAAACAGCTCGGAAATGCTTCTTACGATTATCA	680
DB	774	AGAACTATTAAAGAGAAAGAAACGCAATGGAAACAGCTCGGAAATGCTTCTTACGATTATCA	833
QY	681	CCAGGAAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAAAGATGG	740
DB	834	CCAGGAAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAAAGATGG	893

Qy	741	AGGGCCACAAGGATTTAGCCAAACAACTTTAAGGCCCAATCTAGTCGAAACTAGTTGTCC	800		
Db	894	AGGGCCACAAGGATTTAGCCAAACAACTTTAAGGCCCAATCTAGTCGAAACTAGTTGTCC	953		
Qy	801	CAACATACGGATGGATCCAAATTTATGCCCGCTGATCCAGACTGTGATTTGCTTTATACA	860		
Db	954	CAACATACGGATGGATCCAAATTTATGCCCGCTGATCCAGACTGTGATTTGCTTTATACA	1013		
Qy	861	TAGCAACGATATTTGGATATCTAACATCGTAACACGAGAGAGAGAGACTCACTATGT	920		
Db	1014	TAGCAACGATATTTGGATATCTAACATCGTAACACGAGAGAGAGAGACTCACTATGT	1073		
Qy	921	GCACATGAGCTAGCCAACTGGAAGATGCAATGCACTGAGTGGAGTCCCTATTTGT	980		
Db	1074	GCACATGAGCTAGCCAACTGGAAGATGCAATGCACTGAGTGGAGTCCCTATTTGT	1133		
Qy	981	TCTCCAGAGAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAGCTGAAACAAC	1040		
Db	1134	TCTCCAGAGAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAGCTGAAACAAC	1193		
Qy	1041	TCCAGTGGTGGTAAATTTCTTAAATTTCTATATGAAGAAATGATGAATCTGAGGTGA	1100		
Db	1194	TCCAGTGGTGGTAAATTTCTTAAATTTCTATATGAAGAAATGATGAATCTGAGGTGA	1253		
Qy	1101	AATTAATCATGTTACATCCCTATGTTGGAACAAGAGGGCAGATTCATTCGGTATCC	1160		
Db	1254	AATTAATCATGTTACATCCCTATGTTGGAACAAGAGGGCAGATTCATTCGGTATCC	1313		
Qy	1161	TAAACAGGTAACAGCAATCCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGC	1220		
Db	1314	TAAACAGGTAACAGCAATCCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGC	1373		
Qy	1221	TGAAGAGGATCATAGATGTCATAGATAAGGAACTAAATTCACACCTTTTGAGATTCATTT	1280		
Db	1374	TGAAGAGGATCATAGATGTCATAGATAAGGAACTAAATTCACACCTTTTGAGATTCATTT	1433		
Qy	1281	TGAAGAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAAATATGCTTGTC	1340		
Db	1434	TGAAGAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAAATATGCTTGTC	1493		
Qy	1341	CATCCTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATGCTCACCTGGAATTT	1400		
Db	1494	CATCCTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATGCTCACCTGGAATTT	1553		
Qy	1401	TATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTTGAGTCAGTGCCTGATTC	1460		
Db	1554	TATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTTGAGTCAGTGCCTGATTC	1613		
Qy	1461	TGTGACGCCACTAATTTATCTATGAAGAAACACACACATCTGGATAAATATCCATGACAT	1520		
Db	1614	TGTGACGCCACTAATTTATCTATGAAGAAACACACACATCTGGATAAATATCCATGACAT	1673		
Qy	1521	CTTTCTATGTTTTCCTCCAAAGTCAGAGAGGAAATTTGAGTTTATTTTGGCTCTGATG	1580		
Db	1674	CTTTCTATGTTTTCCTCCAAAGTCAGAGAGGAAATTTGAGTTTATTTTGGCTCTGATG	1733		
Qy	1581	CAAAACAGGTTTCCGTCATTTATACAAATTCATCTATTTTAAAGGAAAGCAATATATA	1640		
Db	1734	CAAAACAGGTTTCCGTCATTTATACAAATTCATCTATTTTAAAGGAAAGCAATATATA	1793		
Qy	1641	ACGATCAGTGTGGGCTGCTGCTCCAGTGAATTTCAAGTGTGCTCTATCAAGAGGAGAT	1700		
Db	1794	ACGATCAGTGTGGGCTGCTGCTCCAGTGAATTTCAAGTGTGCTCTATCAAGAGGAGAT	1853		
Qy	1701	AGCAATTAACAGTGTGAATGGAAATGTTTGGCCGGCATCGATCTAATATCCAAAGTTGA	1760		
Db	1854	AGCAATTAACAGTGTGAATGGAAATGTTTGGCCGGCATCGATCTAATATCCAAAGTTGA	1913		
Qy	1761	TGAAGTCAAGAGCTGGTATATTTTGAAGGACCAAAAGCTCCCTTTTAGAGCATCACCT	1820		
Db	1914	TGAAGTCAAGAGCTGGTATATTTTGAAGGACCAAAAGCTCCCTTTTAGAGCATCACCT	1973		
Qy	1821	GTACGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTC	1880		

```
Db 1974 GTACGTAGTCAAGTACGTTAAATCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTC 2033
Qy 1881 ACATTCCTTGCTGCATCAGTCAGCACTGTGACTCTCTTTATAAGTAAGTATAGTAACAGAA 1940
Db 2034 ACATTCCTTGCTGCATCAGTCAGCACTGTGACTCTCTTTATAAGTAAGTATAGTAACAGAA 2093
Qy 1941 GAATCCACACTGTGTGCTCCCTTTTAAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA 2000
Db 2094 GAATCCACACTGTGTGCTCCCTTTTAAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA 2153
Qy 2001 AACAAAGAAATTTTGGGCGCCACATTTTGGATTTCAGCAGGTCCTCTTCTGACTATACTCC 2060
Db 2154 AACAAAGAAATTTTGGGCGCCACATTTTGGATTTCAGCAGGTCCTCTTCTGACTATACTCC 2213
Qy 2061 TCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA 2120
Db 2214 TCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA 2273
Qy 2121 GCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTGCTGTTTCATATATGGTGCTC 2180
Db 2274 GCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTGCTGTTTCATATATGGTGCTC 2333
Qy 2181 TCAGGTGAGTGTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAAATCCCT 2240
Db 2334 TCAGGTGAGTGTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAAATCCCT 2393
Qy 2241 AGCCTCTCTAGGTTATGTTGTTAGTAGATAGACAAACAGGGGATCTGTCAACGAGGGCT 2300
Db 2394 AGCCTCTCTAGGTTATGTTGTTAGTAGATAGACAAACAGGGGATCTGTCAACGAGGGCT 2453
Qy 2301 TAAATTTGAAGCGGCTTTTAAATATAAAATGG 2332
Db 2454 TAAATTTGAAGCGGCTTTTAAATATAAAATGG 2485

RESULT 20
AR631288
LOCUS AR631288 4685 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 22 from patent US 6844180.
ACCESSION AR631288
VERSION AR631288.1 GI:59770936
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 4685)
  Oi.S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.
  Serine protease genes related to DppIV
  Patent: US 6844180-A 22 18-JAN-2005;
  Ferring BV;;
  NLX;
FEATURES
  Location/Qualifiers
    source 1..4685
    /organism="unknown"
    /mol_type="genomic DNA"
ORIGIN
  Query Match 63.6%; Score 1984; DB 2; Length 4685;
  Best Local Similarity 99.9%; Pred. No. 0;
  Matches 2034; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  Qy 1 AAGTGCTAAAGCCTCGAGGCCAAGCGCGTGTACTGCGCGCGTCTTCTTAGTGCGG 60
  Db 1 AAGTGCTAAAGCCTCGAGGCCAAGCGCGTGTACTGCGCGCGTCTTCTTAGTGCGG 60
  Qy 61 CGTTCCGCGCTGGGTTGTACCGCGCGCCCGCGGAGGACCACTGCAACACGAGCCG 120
  Db 61 CGTTCCGCGCTGGGTTGTACCGCGCGCCCGCGGAGGACCACTGCAACACGAGCCG 120
  Qy 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCATAGCGCACGTGCGGACGG 180
  Db 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCATAGCGCACGTGCGGACGG 180
```

```
Qy 181 TCCGGGCGGGCGGGGGAAGAAATGCAATGCGAGCAGCAATGAAACAGAA 240
Db 181 TCCGGGCGGGCGGGGGAAGAAATGCAATGCGAGCAGCAATGAAACAGAA 240
Qy 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATATTTGAATCAAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATATTTGAATCAAGGATCGG 300
Qy 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATCTCTGAGTCAAGTCTTAAAGCTGCTT 360
Db 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATCTCTGAGTCAAGTCTTAAAGCTGCTT 360
Qy 361 GCCGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCAACCAATGATTTCAATGTTT 420
Db 361 GCCGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCAACCAATGATTTCAATGTTT 420
Qy 421 GTGAAGAGAAATGATCCAGATGGAATCTATTTCAGACAGAAATCTATTACCTTGGCCATGCT 480
Db 421 GTGAAGAGAAATGATCCAGATGGAATCTATTTCAGACAGAAATCTATTACCTTGGCCATGCT 480
Qy 481 GGTGAGACACAGAGAAATACACTGTTTATCTGAAATTTCCAAAACTATCAATAGACGA 540
Db 481 GGTGAGACACAGAGAAATACACTGTTTATCTGAAATTTCCAAAACTATCAATAGACGA 540
Qy 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGTTATCTCGAGAGAGAAATCTATTAAAGAGAAAGAAACGATTTGGAACAGTCCGA 660
Db 601 GGAATGTTATCTCGAGAGAGAAATCTATTAAAGAGAAAGAAACGATTTGGAACAGTCCGA 660
Qy 661 ATTGCTTCTTAAGTATCAACAGAGAGTGAACATTTCTGTTTCAAGCGCGTAGTGGGA 720
Db 661 ATTGCTTCTTAAGTATCAACAGAGAGTGAACATTTCTGTTTCAAGCGCGTAGTGGGA 720
Qy 721 ATTTATCAGTAAAGATGAGGCGGCACAAAGATTTTACCAACACCTTTTAAAGGCCCAAT 780
Db 721 ATTTATCAGTAAAGATGAGGCGGCACAAAGATTTTACCAACACCTTTTAAAGGCCCAAT 780
Qy 781 CTAGTGAGAACTAGTTGTCTCCAAATACATGAGTGAATCCAAATTTATGCGCTGATCCA 840
Db 781 CTAGTGAGAACTAGTTGTCTCCAAATACATGAGTGAATCCAAATTTATGCGCTGATCCA 840
Qy 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGATATCTTAACATGTAACAGAGAA 900
Db 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGATATCTTAACATGTAACAGAGAA 900
Qy 901 GAAAGAGACTCACTTTATGTCACATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGAGACTCACTTTATGTCACATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960
Qy 961 GCTGGAGTCTGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATTTCTTGGCTATTTGGTGG 1020
Db 961 GCTGGAGTCTGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATTTCTTGGCTATTTGGTGG 1020
Qy 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTAATAATTTCTTAGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTAATAATTTCTTAGAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGGAATTTATGTTATCATCCCTATGTTGGAAACAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGGAATTTATGTTATCATCCCTATGTTGGAAACAGGAGG 1140
Qy 1141 GCAGATTTCAATTCGTTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG 1200
Db 1141 GCAGATTTCAATTCGTTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG 1200
Qy 1201 TCAGAAATATGATTTGATGCTGAAGAGGATCATAGATGTCATAGATAGAGAACTAATT 1260
Db 1201 TCAGAAATATGATTTGATGCTGAAGAGGATCATAGATGTCATAGATAGAGAACTAATT 1260
```

```
QY 1261 CAACCTTTTGAGATCTTATTTGAAGGAGTTGCAATATATATGCCCAGAGCTGGAGACTCCT 1320
Db 1261 CAACCTTTTGAGATCTTATTTGAAGGAGTTGCAATATATATGCCCAGAGCTGGAGACTCCT 1320
QY 1321 GAGGGAAATATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTCCAGATAGTG 1380
Db 1321 GAGGGAAATATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTCCAGATAGTG 1380
QY 1381 TTGATCTCAGCTGATATTTATCCAGTAGAGATGATGTTATGAAAAGGAGAGACTC 1440
Db 1381 TTGATCTCAGCTGATATTTATCCAGTAGAGATGATGTTATGAAAAGGAGAGACTC 1440
QY 1441 ATTGAGTCAGTGCCTGATCTCTGTCAGCGCCACTAATATCTATGAAGAAACACAGACATC 1500
Db 1441 ATTGAGTCAGTGCCTGATCTCTGTCAGCGCCACTAATATCTATGAAGAAACACAGACATC 1500
QY 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCACGAAGAGGAATTCGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCACGAAGAGGAATTCGAG 1560
QY 1561 TTTATTTTGGCTCTGAATCAAAACAGAGTTTCCGTCATTTATACAAAATTTACATCTATT 1620
Db 1561 TTTATTTTGGCTCTGAATCAAAACAGAGTTTCCGTCATTTATACAAAATTTACATCTATT 1620
QY 1621 TTAAGGAAAGCAAAATATAAACGATCCAGTGGTGGGCTCGCTCCAGAGTATTCAG 1680
Db 1621 TTAAGGAAAGCAAAATATAAACGATCCAGTGGTGGGCTCGCTCCAGAGTATTCAG 1680
QY 1681 TGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGAGTTCTTGGCCGGCAT 1740
Db 1681 TGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGAGTTCTTGGCCGGCAT 1740
QY 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCCAAAGAC 1800
Db 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCCAAAGAC 1800
QY 1801 TCCCTTTAGAGCATCACCTGTACGTAGTCAGTACGTAAATCTCTGAGAGGTGACAAG 1860
Db 1801 TCCCTTTAGAGCATCACCTGTACGTAGTCAGTAAATCTCTGAGAGGTGACAAG 1860
QY 1861 CTGACTGACCGTGGCTACTCACATCTCTGCTGCATCAGTCAGCTGACATCTGACTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATCTCTGCTGCATCAGTCAGCTGACATCTGACTCTTTATA 1920
QY 1921 AGTAAGTATAGTAACCCAGAGAAATCCACACTGTGTGCTTTCCTTTTACAAGCTATCAAGTCCT 1980
Db 1921 AGTAAGTATAGTAACCCAGAGAAATCCACACTGTGTGCTTTCCTTTTACAAGCTATCAAGTCCT 1980
QY 1981 GRAGATGACCCACTTGCACAAACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAG 2035
Db 1981 GRAGATGACCCACTTGCACAAACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAG 2035
```

```
RESULT 21
AX608745
LOCUS AX608745 4695 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 22 from Patent WO0231134.
ACCESSION AX608745
VERSION AX608745.1 GI:28404308
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1
Qi.S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
Novel serine protease genes related to dppv
Patent: WO 0231134-A 22 18-APR-2002;
Ferring BV (NL)
Location/Qualifiers
1..4695
/organism="Homo sapiens"
```

```
REFERENCE
AUTHORS Qi.S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
TITLE Novel serine protease genes related to dppv
JOURNAL Patent: WO 0231134-A 22 18-APR-2002;
FEATURES
source Ferring BV (NL)
```

```
Query Match 63.6%; Score 1984; DB 2; Length 4695;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2034; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
QY 1 AAGTGTCTAAGCGCTCCGAGGCGCAAGCGCGCTGTCTACTGCGCGCGCTCTTCTTAGTGCCG 60
Db 1 AAGTGTCTAAGCGCTCCGAGGCGCAAGCGCGCTGTCTACTGCGCGCGCTCTTCTTAGTGCCG 60
QY 61 CGTTCGCGCGCTGGGTTGTCAACCAGCGCGCGCGCGCGAGGAGGACCTGCAACACAGGACCG 120
Db 61 CGTTCGCGCGCTGGGTTGTCAACCAGCGCGCGCGCGAGGAGGACCTGCAACACAGGACCG 120
QY 121 GAGTGGAGCGCGCGAGCATGAAAGCGCGCAGCGCCGCTCCATAGGCGCAGCTCGGACCG 180
Db 121 GAGTGGAGCGCGCGAGCATGAAAGCGCGCAGCGCCGCTCCATAGGCGCAGCTCGGACCG 180
QY 181 TCCGCGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAAACAG 240
Db 181 TCCGCGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAAACAG 240
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATGAAATCACAGGATCG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATGAAATCACAGGATCG 300
QY 301 CCTAAATTTGAGCGCTTTTATTTGAGCGGTATTTCTGGAGTCAGCTTAAAGCTGCTT 360
Db 301 CCTAAATTTGAGCGCTTTTATTTGAGCGGTATTTCTGGAGTCAGCTTAAAGCTGCTT 360
QY 361 GCCGATACCAAGAAATATCATGGCTACATGATGGCTTAAGCACCACATGATTTTCTT 420
Db 361 GCCGATACCAAGAAATATCATGGCTACATGATGGCTTAAGCACCACATGATTTTCTT 420
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTCTATTTACCTGCGCATCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTCTATTTACCTGCGCATCT 480
QY 481 GGTGAGAAACAGAGAAATACACTGTTTATTTCTGAAATTCGCAAACTATCAATAGACA 540
Db 481 GGTGAGAAACAGAGAAATACACTGTTTATTTCTGAAATTCGCAAACTATCAATAGACA 540
QY 541 GCAGTCTTAATGCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCAACTGGAATCTAT 600
Db 541 GCAGTCTTAATGCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCAACTGGAATCTAT 600
QY 601 GGAATGTTATTTCTCGAAGAGAACTATTTAAGAGAGAAACCGCATTTGGAACAGTCGGA 660
Db 601 GGAATGTTATTTCTCGAAGAGAACTATTTAAGAGAGAAACCGCATTTGGAACAGTCGGA 660
QY 661 ATTGCTTCTTACGATTTATCCCAAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720
Db 661 ATTGCTTCTTACGATTTATCCCAAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720
QY 721 ATTATACGTTAAAGATGGAGGGCCCAAGAGTTTACGCAACACCTTTTAAGGCCCAAT 780
Db 721 ATTATACGTTAAAGATGGAGGGCCCAAGAGTTTACGCAACACCTTTTAAGGCCCAAT 780
QY 781 CTAGTGGAACCTAGTTGTGCCCAACATACGAGTGGATCCAAATTTATGCCCTGCTGATCCA 840
Db 781 CTAGTGGAACCTAGTTGTGCCCAACATACGAGTGGATCCAAATTTATGCCCTGCTGATCCA 840
QY 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCTGTAACACAGAA 900
Db 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCTGTAACACAGAA 900
QY 901 GAAAGGAGACTCATCTTATGTGCAACATGAGCTAGCCAACTGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCATCTTATGTGCAACATGAGCTAGCCAACTGGAAGAGATGCCAGATCA 960
QY 961 GCTGAGTGCCTTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGCTATTGCTG 1020
```

Db	961		GCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTCCTGGCTATTGGTGG	1020
Qy	1021		TGTCCTCAAAAGCTGAACAACTCCCAAGTGTGTGTAATAATCTTAGAATTTCTATATGAAGAA	1080
Db	1021		TGTCCTCAAAAGCTGAACAACTCCCAAGTGTGTGTAATAATCTTAGAATTTCTATATGAAGAA	1080
Qy	1081		AATGATGAATCTGAGGTGGAATATATCATGTTATCATCCCTATGTTGGAACCAAGGAGG	1140
Db	1081		AATGATGAATCTGAGGTGGAATATATCATGTTATCATCCCTATGTTGGAACCAAGGAGG	1140
Qy	1141		GCAGATTCATTCCTCGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTCACCTTTTAAAGATG	1200
Db	1141		GCAGATTCATTCCTCGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTCACCTTTTAAAGATG	1200
Qy	1201		TCAGAAATAATGATGCTGAAAGGAGATCATAGATGTCATAGATAAGGAACCTAATTT	1260
Db	1201		TCAGAAATAATGATGCTGAAAGGAGATCATAGATGTCATAGATAAGGAACCTAATTT	1260
Qy	1261		CAACCTTTTCAGATTCATTTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGACTCCT	1320
Db	1261		CAACCTTTTCAGATTCATTTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGACTCCT	1320
Qy	1321		GAGGAAATAATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG	1380
Db	1321		GAGGAAATAATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG	1380
Qy	1381		TTGATCTCAGCTGAATATTTATCCAGTAGAGATGATGTTATGGAAGGCGAGAGCTC	1440
Db	1381		TTGATCTCAGCTGAATATTTATCCAGTAGAGATGATGTTATGGAAGGCGAGAGCTC	1440
Qy	1441		ATTGAGTCAGTGCCTGATCTGTGAGGCCACTAATATCTATGAAGAAACCAACGACATC	1500
Db	1441		ATTGAGTCAGTGCCTGATCTGTGAGGCCACTAATATCTATGAAGAAACCAACGACATC	1500
Qy	1501		TGGATAAATATCCAGATCATCTTTCATGTTTTTCCCAAGTCACGAGAGGAATTCAG	1560
Db	1501		TGGATAAATATCCAGATCATCTTTCATGTTTTTCCCAAGTCACGAGAGGAATTCAG	1560
Qy	1561		TTTATTTTTCGCTCGAATCAAAACAGGTTTCCTGTCATTTATACAAAATACATCTATT	1620
Db	1561		TTTATTTTTCGCTCGAATCAAAACAGGTTTCCTGTCATTTATACAAAATACATCTATT	1620
Qy	1621		TTAAAGGAAACCAATATAACGATCAGTGGTGGCTGCCTCCCAAGTATTTCAAG	1680
Db	1621		TTAAAGGAAACCAATATAACGATCAGTGGTGGCTGCCTCCCAAGTATTTCAAG	1680
Qy	1681		TGTCCTATCAAGAGGATAGCAATTCACGTTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Db	1681		TGTCCTATCAAGAGGATAGCAATTCACGTTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Qy	1741		GGATCTAATATCAAGTTGATGAAGTCAGAGGCTGTGTATATTTGAAGGCAACCAAGAC	1800
Db	1741		GGATCTAATATCAAGTTGATGAAGTCAGAGGCTGTGTATATTTGAAGGCAACCAAGAC	1800
Qy	1801		TCCCTTTTAGAGATCATCCTGTAGTACGTTACGTTAATCTCGAGAGGTGACAGG	1860
Db	1801		TCCCTTTTAGAGATCATCCTGTAGTACGTTAATCTCGAGAGGTGACAGG	1860
Qy	1861		CTGACTGACCGTGGCTACTCATTCTTGTGTCATCAGTCAGCTGTGACTGTCTTTTATA	1920
Db	1861		CTGACTGACCGTGGCTACTCATTCTTGTGTCATCAGTCAGCTGTGACTGTCTTTTATA	1920
Qy	1921		AGTAAGTATAGTAAACCAAGAAATCCACATGTGTGTCCTTTTACAGCTATCAAGTCT	1980
Db	1921		AGTAAGTATAGTAAACCAAGAAATCCACATGTGTGTCCTTTTACAGCTATCAAGTCT	1980
Qy	1981		GAGATGACCCCAACTTGCATAACCAAGAAATTTTGGGCCACCACTTTTGGATTTCAG	2035
Db	1981		GAGATGACCCCAACTTGCATAACCAAGAAATTTTGGGCCACCACTTTTGGATTTCAG	2035

AK000290	2778 bp	mRNA	linear	PRI 13-SEP-2003
LOCUS	Homo sapiens cDNA FLJ20283 fis, clone HEP04088.			
DEFINITION				
ACCESSION	AK000290			
VERSION	AK000290.1	GI:7020272		
KEYWORDS	oligo capping; fis (full insert sequence).			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1. Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Ota,T., Nakamura,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Sakurai,Y., Isogai,T. and Sugano,S.			
TITLE	NEDO human cDNA sequencing project			
JOURNAL	Unpublished			
AUTHORS	2 (bases 1 to 2778) Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:flcda@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)			
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).			
FEATURES	Location/Qualifiers			
source	1..2778			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clones="HEP04088"			
	/cell_lines="HepG2"			
	/cell_type="hepatoma"			
	/clone_lib="HEP"			
	/note="Cloning vector pME18SFL3"			
	163..1572			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="BAA91059.1"			
	/db_xref="GI:7020273"			
	/translation="MAAAMETEQLGVIFETADCEINIESQDRPKLEPPFVRYSHSQ LKKLLADTRKYHGYMMAKAPHDPMFVKRNDPDGPHSDRIYYILAMSGENRENTLIFYSEI PKTINRAAVLMSWKPLDLFOATLDYDMYSRBEELRKRKRTGTGVIASYDYHQSNG TELFOAGSGIVHKDGGPQGTQOPLRNLVETSCPNIRMDPKLCPADPWIAFHSN DLWISNITRERBLTYVHNELANMESDASAGVATVLOEFDYSGVWCPCAKETI PSGGKILRIEENDESEVLIHTSPMLSTRADSPRYPKTGTANKPTVTFKMSIEM DAEGRIIDVDRKLIQPFILDFEGVEIKARAGWTPEGKYANSILLDRSOTRLQILVIS PLPLTPVEDDDVMERQRLISVSDSVTPLLIYEETTDIWINIHIDLFHVFPSSHREEIEF IPASECKTGFRHLYKITSI"			
CDS				

Query Match	62.9%	Score 1963;	DB 5;	Length 2778;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 2133;	Conservative 0;	Mismatches 1;	Indels 1;	Gaps 1;
Qy	52	TTAGTGCCTGCTTTCGCGCTTGGGTTGTACCGGCGCGCGCGAGAGCCACTGCAA	111	
Db	1	TTAGTGCCTGCTTTCGCGCTTGGGTTGTACCGGCGCGCGCGAGAGCCACTGCAA	60	
Qy	112	CCAGGACCGAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCAG	171	
Db	61	CCAGGACCGAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCAG	120	
Qy	172	TCCGGACGGTCCGGCGCGCGCGGGAAGGAAATGCAACATGCGCAGCAGCAATGAA	231	

Db	121	TCGGGACGGTCTCGGGCGGGGCGGGGGGAGGAAAAATGCACCATCGCAGCAGCAATGGAA
Qy	232	ACAGAAACAGCTGGGTGTTTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGATCA
Db	181	ACAGAAACAGCTGGGTGTTTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGATCA
Qy	292	CAGGATCGGCCCTAAATTTGGAGCCCTTTTATGTTTGGAGCGGTATCTCTGGAGTCAGCTTAA
Db	241	CAGGATCGGCCCTAAATTTGGAGCCCTTTTATGTTTGGAGCGGTATCTCTGGAGTCAGCTTAA
Qy	352	AAGCTGCTTGGCGGATACACAGAAATATCATGGCTACATGATGGCTTAAGGCAACCATGAT
Db	301	AAGCTGCTTGGCGGATACACAGAAATATCATGGCTACATGATGGCTTAAGGCAACCATGAT
Qy	412	TTTCATGTTTGTGAAGAGGAATGATCCAGATGGAACCTCATTCAGACAGATCTTATTACCTT
Db	361	TTTCATGTTTGTGAAGAGGAATGATCCAGATGGAACCTCATTCAGACAGATCTTATTACCTT
Qy	472	GCCATGTCGTGTAGACAGAGAAATACACTGTTTTHATCTGAAATTCCTCAAAACTATC
Db	421	GCCATGTCGTGTAGACAGAGAAATACACTGTTTTHATCTGAAATTCCTCAAAACTATC
Qy	532	AATAGACGACGAGCTTATGCTCTCTCTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACA
Db	481	AATAGACGACGAGCTTATGCTCTCTCTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACA
Qy	592	CTGGACTATCGAATGTATCTCTCGAGAGAGAACTATTAAGAGAAAGAAAAAGCAATGGGA
Db	541	CTGGACTATCGAATGTATCTCTCGAGAGAGAACTATTAAGAGAAAGAAAAAGCAATGGGA
Qy	652	ACAGTCGGAATTCGCTTCTTACGATATCACCAAGGAAGTGGAACTATTCGTTTCAAGCC
Db	601	ACAGTCGGAATTCGCTTCTTACGATATCACCAAGGAAGTGGAACTATTCGTTTCAAGCC
Qy	712	GGTAGTGGAAATTTATCACGTAAGAAAGTGGAGGGGCCAAGAGATTTACGCAACCAACTTTTA
Db	661	GGTAGTGGAAATTTATCACGTAAGAAAGTGGAGGGGCCAAGAGATTTACGCAACCAACTTTTA
Qy	772	AGGCCCAATCTAGTGGAAACTAGTTGTGCCCAACATACGGATGGATCCAAATTTATGCCCC
Db	721	AGGCCCAATCTAGTGGAAACTAGTTGTGCCCAACATACGGATGGATCCAAATTTATGCCCC
Qy	832	GCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAAACATGTA
Db	781	GCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAAACATGTA
Qy	892	ACCAGAGAAAGAAAGAGAGACTCACATATGTGCAACATGAGCTAGCCCAACATCGAGAGAT
Db	841	ACCAGAGAAAGAAAGAGAGACTCACATATGTGCAACATGAGCTAGCCCAACATCGAGAGAT
Qy	952	GCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAGATTTGATAGATATCTGGC
Db	901	GCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAGATTTGATAGATATCTGGC
Qy	1012	TATTTGTTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTAGAAATCTTA
Db	961	TATTTGTTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTAGAAATCTTA
Qy	1072	TATGAAGAAAAATGATGAATCTGAGTGGGAATTTATTCATGTATCATGCCCTATGTTGGAA
Db	1021	TATGAAGAAAAATGATGAATCTGAGTGGGAATTTATTCATGTATCATGCCCTATGTTGGAA
Qy	1132	ACAAGAGGCGAGATTCATTTCCGTTATCTTAAACAGGTACAGCAAAATCTCAAGATCACT
Db	1081	ACAAGAGGCGAGATTCATTTCCGTTATCTTAAACAGGTACAGCAAAATCTCAAGATCACT
Qy	1192	TTTAAAGATGTCAGAAATTAATGATGTGCTGAAGGAAGGATCATAGATGTCATAGATAAG
Db	1141	TTTAAAGATGTCAGAAATTAATGATGTGCTGAAGGAAGGATCATAGATGTCATAGATAAG
Qy	1252	GNACTAATTCAACTTTTGAGATTCCTATTTGAAGGAGTTGAATATATTTGCGAGAGCTGGA
Db	1201	GNACTAATTCAACTTTTGAGATTCCTATTTGAAGGAGTTGAATATATTTGCGAGAGCTGGA

QY	1312	TGGA	CTCTCTGAGGGGAAAATATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCCTA	1371
Db	1261	TGGA	CTCTCTGAGGGGAAAATATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCCTA	1320
QY	1372	CAGATAGT	GTTTGATCTCACTCGAAATATTTATCCAGTAGAAGATGATGTTATCGGAAGG	1431
Db	1321	CAGATAGT	GTTTGATCTCACTCGAAATATTTATCCAGTAGAAGATGATGTTATCGGAAGG	1380
QY	1432	CAGAGACT	CAITTAGTCACTGCTTGATTCGTGTGACGCCACTAATATCTATGAAGAACCA	1491
Db	1381	CAGAGACT	CAITTAGTCACTGCTTGATTCGTGTGACGCCACTAATATCTATGAAGAACCA	1440
QY	1492	ACAGACAT	CTCGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGATCGAAGAG	1551
Db	1441	ACAGACAT	CTCGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGATCGAAGAG	1500
QY	1552	GAATTCGAGTTTATTTTTTGCTCTGATGCAAAACAGGTTTCCGTCTATTTATACAAAT	1611	
Db	1501	GAATTCGAGTTTATTTTTTGCTCTGATGCAAAACAGGTTTCCGTCTATTTATACAAAT	1560	
QY	1612	ACATCTA	TTTTAAGGAAAGCAATATAAACGATCCAGTGGTGGGTGCTGCTCCAAAGT	1671
Db	1561	ACATCTA	TTTTAAGGAAAGCAATATAAACGATCCAGTGGTGGGTGCTGCTCCAAAGT	1619
QY	1672	GATTTCAAGTGTCTTATCAAAGAGGATAGCAAAATACCAAGTGGTGAATGGGAAGTTC	1731	
Db	1620	GATTTCAAGTGTCTTATCAAAGAGGAGATAGCAAAATACCAAGTGGTGAATGGGAAGTTC	1679	
QY	1732	GGCCGGCATCGATCTAATATCCAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGC	1791	
Db	1680	GGCCGGCATCGATCTAATATCCAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGC	1739	
QY	1792	ACCAAGACTCCCCTTTTAGAGCATCACTGTACGTAGTACGTAAATCCTCGAGAG	1851	
Db	1740	ACCAAGACTCCCCTTTTAGAGCATCACTGTACGTAGTACGTAAATCCTCGAGAG	1799	
QY	1852	GTGACAAGGCTGACTCAACCGTGGCTACTCACATCTCTGTCGATCAGTCAGCATGTGC	1911	
Db	1800	GTGACAAGGCTGACTCAACCGTGGCTACTCACATCTCTGTCGATCAGTCAGCATGTGC	1859	
QY	1912	TTCTTTTATAAGTAAATAGTAAACAGAGAGATCCACACTGTGTCTCCCTTTACAAAGCTA	1971	
Db	1860	TTCTTTTATAAGTAAATAGTAAACAGAGAGATCCACACTGTGTCTCCCTTTACAAAGCTA	1919	
QY	1972	TCAAGTCCTGAGATGACCCACTTTGCAAAACAAAGGAAATTTGGGCCACCATTTTGGAT	2031	
Db	1920	TCAAGTCCTGAGATGACCCACTTTGCAAAACAAAGGAAATTTGGGCCACCATTTTGGAT	1979	
QY	2032	TCAGCAGGTCCTTCTCGACTATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACT	2091	
Db	1980	TCAGCAGGTCCTTCTCGACTATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACT	2039	
QY	2092	GGATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATAT	2151	
Db	2040	GGATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATAT	2099	
QY	2152	CCTACTGTGCTTTCATATATGGTGGTCCCTCAGGT	2186	
Db	2100	CCTACTGTGCTTTCATATATGGTGGTCCCTCAGGT	2134	

RESULT 23
BD157001
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BD157001 2161 bp DNA linear PAT 17-JAN-2003
Primer for synthesizing full-length cDNA and use thereof.

DEFINITION	Homo sapiens cDNA FLJ14920 fis, clone PLACE1007416, weakly similar to D1PEP1DYL PEPTIDASE IV (BC 3.4.14.5).
ACCESSION	AK027826
VERSION	AK027826.1 GI:14042789
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirao, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Hosoda, K., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T. O., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arima, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, P., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigetake, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
TITLE	Complete sequencing and characterization of 21,243 full-length human cDNAs
JOURNAL	Nat. Genet. 36 (1), 40-45 (2004)
PUBMED	14702039
REFERENCE	2
AUTHORS	Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 2161)
AUTHORS	Isogai, T. and Otsuki, T.
TITLE	Direct Submission
JOURNAL	Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: flj-cdnasnifty.com, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES	Location/Qualifiers
source	1..2161
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="PLACE1007416"

CDS	/tissue type="placenta"
	/clone_lib="PLACE1"
	/note="Cloning vector: pME18SFL3"
	16..1914
	/code="unnamed protein product"
	/codon_start=1
	/protein_id="BAB55395.1"
	/db_xref="GI:14042790"
	/translation="MDPKLCPADPDMDIAFTHSDNIWISNIVTREERLTVNHELANM
	REDARSAGVATFVLQESFDYSGVWCPKAEITTPSGKILRILYBENDESEVEIHVT
	SEMLERADRSFRYPKGTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIQFELFPGV
	EYIARAGWTPEGVKMSILLDRSQTSLQIVLISPELFIPEVDDVMERQRLIESVPSV
	TPLIITYETTDIWINIHDIHFVFPQSHEEEIEFIFASECTGRHLIKYKLSILKESKY
	KRSSGLVPAPDFKCPKEETAITSGEWEVLGRHSNIQVDEVRLLVYFEGTKUSPLE
	HLHYVSNVNPCEVTRLTDSYSHSCISQHCDFFIKSYKNQKPHCVLSLKUSPED
	DPYCTKEFWATILDSAGPLPDYTPPEIFSFPESTGFTLYGLMKPHDLPQGGKYPTV
	LFYGGPQVQVNNRFKGVKVFRLNTLASLYVVVVDNRGSHRGKLPFGAPKYKXV
	AIAGAPVTIMIFDYGTYERNYGHDPQNEQGYLVGSVAMQAEKFPSPENRLLHLHGL
	DENVHFAHTSILLSLFLVRAGKPYDLQYFQBERHSIRVPESGEHVELHLHLVQLNLS
	RTAALKVI"
ORIGIN	
Query Match	46.5%; Score 1450; DB 5; Length 2161;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 1500; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	832 GCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATTGATATCTAACATCGTA 891
DB	
QY	37 GCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATTGATATCTAACATCGTA 96
DB	
QY	892 ACCAGAGAAGAGGAGACTCTATTGTGTCACATAGTAGCCCAACATCGAAGAAGAT 951
DB	
QY	97 ACCAGAGAAGAGGAGACTCTATTGTGTCACATAGTAGCCCAACATCGAAGAAGAT 156
DB	
QY	952 GCCAGATCAGCTGAGTGCCTACTCTTGTCTTCCAAAGAGAAATTTGATAGATATTTCTGGC 1011
DB	
QY	157 GCCAGATCAGCTGAGTGCCTACTCTTGTCTTCCAAAGAGAAATTTGATAGATATTTCTGGC 216
DB	
QY	1012 TATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTA 1071
DB	
QY	217 TATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTA 276
DB	
QY	1072 TATGAAGAAATATGATGAATCTGAGGTGGAAATATTATTCATGTTTACATCCCTATGTTGAA 1131
DB	
QY	277 TATGAAGAAATATGATGAATCTGAGGTGGAAATATTATTCATGTTTACATCCCTATGTTGAA 336
DB	
QY	1132 ACAAGGAGGCGAGATTCATTCCGTTATCTCTTAAACAGGTACAGCAATCTTAAAGTCACT 1191
DB	
QY	337 ACAAGGAGGCGAGATTCATTCCGTTATCTCTTAAACAGGTACAGCAATCTTAAAGTCACT 396
DB	
QY	1192 TTTAAGATGTCAGAAATAATGATGATGCTGGAAGGAGGATCATAGATGTCATAGATAAG 1251
DB	
QY	397 TTTAAGATGTCAGAAATAATGATGATGCTGGAAGGAGGATCATAGATGTCATAGATAAG 456
DB	
QY	1252 GAACTAAATCAACCTTTTGAGATTCTATTTTGAAGGAGTTGAATATATATGCGAGAGCTGA 1311
DB	
QY	457 GAACTAAATCAACCTTTTGAGATTCTATTTTGAAGGAGTTGAATATATATGCGAGAGCTGA 516
DB	
QY	1312 TGGACTCTCTGAGGGAATAATATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCCGCTTA 1371
DB	
QY	517 TGGACTCTCTGAGGGAATAATATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCCGCTTA 576
DB	
QY	1372 CAGATAGTGTGATCTCACCTGGAATTTATTTCCAGTAGAAGATGATGTTATGGAAGG 1431
DB	
QY	577 CAGATAGTGTGATCTCACCTGGAATTTATTTCCAGTAGAAGATGATGTTATGGAAGG 636
DB	
QY	1432 CAGAGACTCATTTGAGTGCAGTGCCTGATTTCTGTGAGCCCATTAATTATCTATGAGAAACA 1491
DB	
QY	637 CAGAGACTCATTTGAGTGCAGTGCCTGATTTCTGTGAGCCCATTAATTATCTATGAGAAACA 696
DB	
QY	1492 ACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAG 1551
DB	
QY	697 ACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAG 756
DB	

QY	1552	GAATTTGAGTTATTTTTCCTCTGAATGCAAAACAGGTTTCGTCAATTTATACAAAATT	1611
Db	757	GAATTTGAGTTATTTTTCCTCTGAATGCAAAACAGGTTTCGTCAATTTATACAAAATT	816
QY	1612	ACATCTATTTTAAAGGAAGCAATATAAACGATCAAGTGGTGGCTGCTCTCAAGT	1671
Db	817	ACATCTATTTTAAAGGAAGCAATATAAACGATCAAGTGGTGGCTGCTCTCAAGT	876
QY	1672	GATTTCAAGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTT	1731
Db	877	GATTTCAAGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTT	936
QY	1732	GGCGGCAATGATCTAATATATCAAGTTGATGAAGTCAAGAGCTGGTATATTTGAAGGC	1791
Db	937	GGCGGCAATGATCTAATATATCAAGTTGATGAAGTCAAGAGCTGGTATATTTGAAGGC	996
QY	1792	ACCAAGACTCCCTTTTAGAGCATCACTGTAGTACAGTACGTTAAATCCCTGGAGAG	1851
Db	997	ACCAAGACTCCCTTTTAGAGCATCACTGTAGTACAGTACGTTAAATCCCTGGAGAG	1056
QY	1852	GTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACTGTGAC	1911
Db	1057	GTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACTGTGAC	1116
QY	1912	TTCTTTTATAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTA	1971
Db	1117	TTCTTTTATAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTA	1176
QY	1972	TCAAGTCTTGAAGATGACCCCACTTGCAAAACAAAAGGAATTTTGGGCCACCATTTGGAT	2031
Db	1177	TCAAGTCTTGAAGATGACCCCACTTGCAAAACAAAAGGAATTTTGGGCCACCATTTGGAT	1236
QY	2032	TCAGCAGGTCCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACT	2091
Db	1237	TCAGCAGGTCCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACT	1296
QY	2092	GGATTTACATTTGATGGGATGCTTACAGGCTCATGATCTACAGCTGGAAGAAATAT	2151
Db	1297	GGATTTACATTTGATGGGATGCTTACAGGCTCATGATCTACAGCTGGAAGAAATAT	1356
QY	2152	CCTACTGTCTGTTATATATGTTGCTCTCAGTGCAAGTGGTGGTGAATATCGGTTTAA	2211
Db	1357	CCTACTGTCTGTTATATATGTTGCTCTCAGTGCAAGTGGTGGTGAATATCGGTTTAA	1416
QY	2212	GGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTTAGGTTATGTTGTAGTGATA	2271
Db	1417	GGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTTAGGTTATGTTGTAGTGATA	1476
QY	2272	GACACAGGGGATCTCTGTCAACGAGGCTTAAATTTGAAGCGCGCTTTAAATATAAAATG	2331
Db	1477	GACACAGGGGATCTCTGTCAACGAGGCTTAAATTTGAAGCGCGCTTTAAATATAAAATG	1536
QY	2332	G 2332	
Db	1537	G 1537	

RESULT 26
AX405771
LOCUS
DEFINITION
Sequence 186 from Patent WO0222660.
ACCESSION
AX405771.1 GI:21438981
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,
Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. F.

TITLE	Novel nucleic acids and polypeptides									
JOURNAL	Patent: WO 0222660-A 186 21-MAR-2002;									
FEATURES	HYSEQ, INC. (US)									
source	Location/Qualifiers									
CDS	1..2668 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 234..2408 /note="unnamed protein product" /codon_start=1 /protein_id="CAD34878.1" /db_xref="GI:21438982" /translation="MAAAMETQLGVEIFETADCEENIESQDRPKLEPFYVRYMSQ LKKLLAATRYKHGMKMAKPHDFMFKRNDPDPHSDRIYVLLAMGENRENTLFTSEI PKTINRAAVLMSKPLLDLUFQOQPLRPNLIVETSCNIRMDPKLCPADDMTAFTHSN DIMLSNIVTRERRLTYVHNELANBEDARSAGVATFVLOEBFDYSGYVWCPKAEIT PSGKILRLIYENDSEVEIIHVTSPMLETRADSFYKQFDTANPKVTFKMSIMI DAGRIIDVIDKELIQPFLFEGVEIARAGATPEGYKANGILLDRSQTRLIQVILIS FELFIPVEDDVMERQRLIESVPDSVPLIYYEETDIWINIHDIHFVFPQSHHEEIEF IFASECKTGFRHLKYKITSILKSKYKRSGGLPAPDFKCPILKEBIAITSGBEVIGR HGSNIQVDEVRLLYFEGTKDSPLHLLYVSVVNFGEVRLTRDGYSHSCISQCHD FFISKYNQNPCHVSLYKLSPEDDPTCKTFEWAATILDSAGPLDPTVPEIFSGFHS TTGTLYGMLYKPHDLQPGKKYPTLVFIYGGPQVAIAGAPVLMIFYDTGYTERVMGH PDQNEQSYLGSVMAQAEKPPSEPNRLLLHLGHFLDENVHFAHTSILLSFLVRAGPKPYD LQIYPOERHSIRVPSEGEHYELHLLHQENLGSRIAALKVI"									
ORIGIN	44.1%; Score 1377; DB 2; Length 2668; Query Match Best Local Similarity 99.9%; Pred. No. 0; Matches 1427; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	759	GCAACAAACCTTTAAGGCCCAATCTAGTGGAAAATAGTTGTGCCAACATACGGATGGATCC	818							
Db	605	GCAACAAACCTTTAAGGCCCAATCTAGTGGAAAATAGTTGTGCCAACATACGGATGGATCC	664							
QY	819	AAAATTATGCCCGCTGATCCAGACTGGAATTCCTTTTATACATAGCAACGATATTTGGAT	878							
Db	665	AAAATTATGCCCGCTGATCCAGACTGGAATTCCTTTTATACATAGCAACGATATTTGGAT	724							
QY	879	ATCTAAACATCGTAAACAGAGAGAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAA	938							
Db	725	ATCTAAACATCGTAAACAGAGAGAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAA	784							
QY	939	CATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAAGATTTGA	998							
Db	785	CATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAAGATTTGA	844							
QY	999	TAGATATTCGTGCTATTGCTGTGCCAAAGCTGAAAACAACCTCCAGTGGTGGTAAAAAT	1058							
Db	845	TAGATATTCGTGCTATTGCTGTGCCAAAGCTGAAAACAACCTCCAGTGGTGGTAAAAAT	904							
QY	1059	TCTTAGAATCTATATGAAGAAAATGATGAATCTGAGTGGGAATTTATCATGTTACATC	1118							
Db	905	TCTTAGAATCTATATGAAGAAAATGATGAATCTGAGTGGGAATTTATCATGTTACATC	964							
QY	1119	CCCTATGTTGGAACAAGAGGGCAGATTCATTCCGTTATCTCTAAACACAGGTACAGCAA	1178							
Db	965	CCCTATGTTGGAACAAGAGGGCAGATTCATTCCGTTATCTCTAAACACAGGTACAGCAA	1024							
QY	1179	TCTTAAAGTCATTTTAAAGATGTGAGAATAATGATGATGCTGAAGGAAGATCATAGA	1238							
Db	1025	TCTTAAAGTCATTTTAAAGATGTGAGAATAATGATGATGCTGAAGGAAGATCATAGA	1084							
QY	1239	TGTCATAGATAGGAACCTAATTCACCTTTTGGAGATCTATTGGAAGGAGTTCGAATAT	1298							
Db	1085	TGTCATAGATAGGAACCTAATTCACCTTTTGGAGATCTATTGGAAGGAGTTCGAATAT	1144							
QY	1299	TGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCATCTCTACTAGATCGCTC	1358							
Db	1145	TGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCATCTCTACTAGATCGCTC	1204							
QY	1359	CCAGACTCGCCTACAGATAGTGTGTGATCTCTCACTGAATTTATTTATCCCAAGTAGAAGATGA	1418							

source	1.. 4309																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						</
--------	----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

```
QY 2708 TAGTGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCAGAGAGACACAGCATAA 2767
D 2190 TAGTGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCAGAGAGACACAGCATAA 2249
QY 2768 GAGTTTCCGTAATCGGGAGAACATTTATGACATGCTTTTGGCTACCTTCAAGAAACC 2827
D 2250 GAGTTTCCGTAATCGGGAGAACATTTATGACATGCTTTTGGCTACCTTCAAGAAACC 2309
QY 2828 TTGGATCAGGTAATGCTCTCTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGG 2887
D 2310 TTGGATCAGGTAATGCTCTCTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGG 2369
QY 2888 TATACACTGGCTATTATACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATGATC 2947
D 2370 TATACACTGGCTATTATACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATGATC 2429
QY 2948 ATACATTTTGATACCTGCCATGPAACATCTACTCTCGAAATAAATGTGGTGCATGCA 3007
D 2430 ATACATTTTGATACCTGCCATGPAACATCTACTCTCGAAATAAATGTGGTGCATGCA 2489
QY 3008 GGGGTCTACGGTTTGTGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAATG 3067
D 2490 GGGGTCTACGGTTTGTGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAATG 2549
QY 3068 ATACATTTCTCGAGAGACCCAGCAATACCATAAGAAATTTACTAAAAAABAAA 3120
D 2550 ATACATTTCTCGAGAGACCCAGCAATACCATAAGAAATTTACTAAAAAABAAA 2602

RESULT 28
AX608737
LOCUS AX608737 4309 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 14 from Patent WO0231134.
ACCESSION AX608737
VERSION AX608737.1 GI:28404304
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE 1
AUTHORS Qi, S., Akimanya, K.O., Riviere, P.J. and Junlen, J.L.
TITLE Novel serine protease genes related to dppv
JOURNAL Patent: WO 0231134-A 14 18-APR-2002;
Ferring BV (NL)
FEATURES
Source Location/Qualifiers
1..4309
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 38.2%; Score 1191; DB 2; Length 4309;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1371; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1750 ATCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTA 1809
D 1230 ATCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTA 1289
QY 1810 GAGCATCACCTGACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAGGCTGACTGAC 1869
D 1290 GAGCATCACCTGACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAGGCTGACTGAC 1349
QY 1870 CGTGGCTACTACATTTCTTGCTGATCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 1929
D 1350 CGTGGCTACTACATTTCTTGCTGATCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 1409
QY 1930 AGTAACCAAGAAATCCACATCTGTGTCCTTTTACAGATCAAGTCTCTGAGATGAC 1989
D 1410 AGTAACCAAGAAATCCACATCTGTGTCCTTTTACAGATCAAGTCTCTGAGATGAC 1469
```

```
QY 1990 CCAACTTGCAAAACAAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCCTCTTCCT 2049
D 1470 CCAACTTGCAAAACAAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCCTCTTCCT 1529
QY 2050 GACTATACCTCTCCAGAAATTTTCTCTTTTCAAAAGTACTACTGATTTACATTTGATGGG 2109
D 1530 GACTATACCTCTCCAGAAATTTTCTCTTTTCAAAAGTACTACTGATTTACATTTGATGGG 1589
QY 2110 ATGCTCTCAAGSCCTCATGATCTACAGCTCGAAAGAAATATCTACTGCTGCTTTCATA 2169
D 1590 ATGCTCTCAAGSCCTCATGATCTACAGCTCGAAAGAAATATCTACTGCTGCTTTCATA 1649
QY 2170 TATGTGG--TCCTCAGGTCAGTTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCC 2227
D 1650 TATGTGGTCTCTCAGGTCAGTTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCC 1709
QY 2228 GCTTCGAATACCTAGCTCTCTAGGTTATGTGGTTAGTGATAGACAAACAGGGGATCCT 2287
D 1710 GCTTCGAATACCTAGCTCTCTAGGTTATGTGGTTAGTGATAGACAAACAGGGGATCCT 1769
QY 2288 GTCCAGGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAAATAGAAATTTG 2347
D 1770 GTCCAGGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAAATAGAAATTTG 1829
QY 2348 AGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTTAGATC 2407
D 1830 AGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTTAGATC 1889
QY 2408 GTGTGGGCATCCAGCGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGA 2467
D 1890 GTGTGGGCATCCAGCGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGA 1949
QY 2468 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATTTCTATG 2527
D 1950 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATTTCTATG 2009
QY 2528 ATACAGATATACAGGAAAGCTTATATGGTCCACCTGACAGAGATGAACAGGCTATTAAT 2587
D 2010 ATACAGATATACAGGAAAGCTTATATGGTCCACCTGACAGAGATGAACAGGCTATTAAT 2069
QY 2588 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATGGTTTACTGCTCT 2647
D 2070 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATGGTTTACTGCTCT 2129
QY 2648 TACATGTTTCTCTGGATGAGATGTCATTTTGCAATACCATGATATATCTAGTATTTT 2707
D 2130 TACATGTTTCTCTGGATGAGATGTCATTTTGCAATACCATGATATATCTAGTATTTT 2189
QY 2708 TAGTGAGGCTCGAAAGCCATATGATTTTACAGATCTATCTCAGAGAGACACAGCATAA 2767
D 2190 TAGTGAGGCTCGAAAGCCATATGATTTTACAGATCTATCTCAGAGAGACACAGCATAA 2249
QY 2768 GAGTTTCTGAAATCGGAGAACATTTATGAACTGCTTTTGGACTTCTTCAAGAAACC 2827
D 2250 GAGTTTCTGAAATCGGAGAACATTTATGAACTGCTTTTGGACTTCTTCAAGAAACC 2309
QY 2828 TTGGATCAGGTAATGCTCTCTTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGG 2887
D 2310 TTGGATCAGGTAATGCTCTCTTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGG 2369
QY 2888 TATACACTGGCTATTATACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATGATC 2947
D 2370 TATACACTGGCTATTATACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATGATC 2429
QY 2948 ATCAATTTTGATACCTGCTGATGTAACATCTACTCTCGAAATAAATGTGGTGCATGCA 3007
D 2430 ATCAATTTTGATACCTGCTGATGTAACATCTACTCTCGAAATAAATGTGGTGCATGCA 2489
QY 3008 GGGGTCTACGGTTTGTGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAATG 3067
D 2490 GGGGTCTACGGTTTGTGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAATG 2549
QY 3068 ATACATATTTCTCGAGAGACCCAGCAATACCATAAGAAATTTACTAAAAAABAAA 3120
```

[illegible]

source		1. .1669	/organism="unknown"		/mol_type="genomic DNA"	
Query Match		32.8%;	Score 1023;	DB 2;	Length 1669;	
Best Local Similarity		100.0%;	Pred. No. 0;			
Matches 1023;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
ORIGIN						
QY	1164	AACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGA	1223			
DB	1	AACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGA	60			
QY	1224	AGGAAGGATCATGATGTCATAGATAAGGAACATAATCAACCTTTGAGATTCATTATGA	1283			
DB	61	AGGAAGGATCATGATGTCATAGATAAGGAACATAATCAACCTTTGAGATTCATTATGA	120			
QY	1284	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGAAAAATATGCTGGTCCAT	1343			
DB	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGAAAAATATGCTGGTCCAT	180			
QY	1344	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATATTATAT	1403			
DB	181	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATATTATAT	240			
QY	1404	CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGATCTGT	1463			
DB	241	CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGATCTGT	300			
QY	1464	GAGCCCACTAATATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGATCTGT	1523			
DB	301	GAGCCCACTAATATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGATCTGT	360			
QY	1524	TCATGTTTTTCCCAAGTACGAAGAGGAATTTGAGTTATTTTTCCTTCGATGCAAA	1583			
DB	361	TCATGTTTTTCCCAAGTACGAAGAGGAATTTGAGTTATTTTTCCTTCGATGCAAA	420			
QY	1584	AACAGGTTTCCGTCATTTATACAAATTTACATCTTTTAAAGAAAGCAATATAAGC	1643			
DB	421	AACAGGTTTCCGTCATTTATACAAATTTACATCTTTTAAAGAAAGCAATATAAGC	480			
QY	1644	ATCCAGTGTGGCTGCTGCTCCAGTGTATTCAGTGATTTCAAGTGCCTATCAAGAGAGATAGC	1703			
DB	481	ATCCAGTGTGGCTGCTGCTCCAGTGTATTCAGTGATTTCAAGTGCCTATCAAGAGAGATAGC	540			
QY	1704	AATACCAAGTGTGAATCGGAAGTTCTTGGCCGCGATGGATCTAAATATCAAGTTGATGA	1763			
DB	541	AATACCAAGTGTGAATCGGAAGTTCTTGGCCGCGATGGATCTAAATATCAAGTTGATGA	600			
QY	1764	AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACTGTA	1823			
DB	601	AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACTGTA	660			
QY	1824	CGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA	1883			
DB	661	CGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA	720			
QY	1884	TTCTTGTGCAATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAA	1943			
DB	721	TTCTTGTGCAATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAA	780			
QY	1944	TCCACACTGTGTGCTTCTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAC	2003			
DB	781	TCCACACTGTGTGCTTCTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAC	840			
QY	2004	AAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTCTGACTACTACTCTCC	2063			
DB	841	AAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTCTGACTACTACTCTCC	900			
QY	2064	AGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGCC	2123			
DB	901	AGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGCC	960			

QY	2124	TCATGATCTACAGCTCGAAGAAATATCTACTGCTGCTTTCATATATGTTGGTCTCTCA	2183
DB	961	TCATGATCTACAGCTCGAAGAAATATCTACTGCTGCTTTCATATATGTTGGTCTCTCA	1020
QY	2184	GGT 2186	
DB	1021	GGT 1023	

RESULT 31	AF221636	1669 bp	mRNA	linear	PRI 05-NOV-2000
LOCUS	Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, partial cds,				
DEFINITION	alternatively spliced.				
ACCESSION	AF221636				
VERSION	AF221636.1	GI:11095191			
KEYWORDS					
SOURCES	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
	Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1669)				
AUTHORS	Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W. and Gorrell,M.D.				
TITLE	Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8				
JOURNAL	Eur. J. Biochem. 267 (20), 6140-6150 (2000)				
PUBMED	11012666				
REFERENCE	2 (bases 1 to 1669)				
AUTHORS	Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.				
TITLE	Direct Submission				
JOURNAL	Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver Centre, Centenary Institute of Cell Biology and Cancer Medicine, Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia				
FEATURES	Location/Qualifiers				
source	1. .1669				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="15"				
	/map="15q22"				
	/clone="T8"				
	/tissue_type="placenta"				
	/clone_lib="lambdaTripleExTM: Clontech"				
gene	<1. 1669				
CDS	<1 1399				
	/gene="DPP8"				
	/gene="DPP8"				
	/note="serine protease; member of prolyl oligopeptidase S9b family belonging to enzyme clan SC; isoform missing the catalytic serine required for enzyme activity; additional splice variants of this mRNA are found in GenBank Accession Numbers AF221634, AF221635 and AF221637"				
	/codon_start=2				
	/product="dipeptidyl peptidase 8"				
	/protein_id="AAG29768.1"				
	/db_xref="GI:11095192"				
	/translation="TGTANPKVTFKMSIMIDAEGRIIDVIDKELIQPEILFEGVEY IARAWTEPKRYAWSILLDRSQRLQIVLISPELFIPEVDDVMERQRLIESVPDSVTP LIIVETDWINIHDFVFPQSHEEIEFIPASECKTGFRHLVYKITSILKESKYKR SSGSLPAPDFKCPKEEIAITGEWEVLRGHSNIQVDEVRLLVYFEGTKDSPLSHH LYVSYVNPGEVRLTRDYGSHSCCSQCHDFISKYQKNPHCVLYKLSSPDDP TCKTKEFWATILDSAGLPDYTPTEIFSFESTFTLYGMLYKPHDLQPKKYPTVLF IYGPQVAIAGAVTLMVIFDYTERYMGHPDQEQYGLGSAVMAQKPFSEPNRL LLHGFDENVHFAVTSILLSFLVLRAGKPYDLQIYQPERHSIRVPESGHEVELHLHY LQENLGSRIAALKVI"				

ORIGIN

Query Match	32.8%;	Score 1023;	DB 5;	Length 1669;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1023;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1164	AA	CAGGTACGACCAATCCCTAAAGTGCACATTTTAAAGTGCAGAAATTAATGATTGATGCTGA	1223
Db	1	AR	CAGGTACGACCAATCCCTAAAGTGCACATTTTAAAGTGCAGAAATTAATGATTGATGCTGA	60
QY	1224	AG	GAAGGATCATAGATGTCATAGATAAGAACTAAATTAACCTTTTGAGATTCATTGGA	1283
Db	61	AG	GAAGGATCATAGATGTCATAGATAAGAACTAAATTAACCTTTTGAGATTCATTGGA	120
QY	1284	AG	GATTTGAATATATTCAGAGCTGGATGACCTCTGAGGAAAAATATCTTTGGTCCAT	1343
Db	121	AG	GATTTGAATATATTCAGAGCTGGATGACCTCTGAGGAAAAATATCTTTGGTCCAT	180
QY	1344	CT	ACTAGATCGCTCCAGACTCGCTACAGATAGTTGATCTCACCTGAATATTAT	1403
Db	181	CT	ACTAGATCGCTCCAGACTCGCTACAGATAGTTGATCTCACCTGAATATTAT	240
QY	1404	CC	CAGTGAAGATGATGTTATGGAAGGAGAGACTCAATGAGTCAGTGGCTGATCTGT	1463
Db	241	CC	CAGTGAAGATGATGTTATGGAAGGAGAGACTCAATGAGTCAGTGGCTGATCTGT	300
QY	1464	GA	CCCACTAATATCTATCAAGAAAAACACAGACATCTGGATAAATATCCATGACATCTT	1523
Db	301	GA	CCCACTAATATCTATCAAGAAAAACACAGACATCTGGATAAATATCCATGACATCTT	360
QY	1524	TC	ATGTTTTTCCCAAGTCCAGAGGAATATGATTTATTTTGGCTGGAATGCAA	1583
Db	361	TC	ATGTTTTTCCCAAGTCCAGAGGAATATGATTTATTTTGGCTGGAATGCAA	420
QY	1584	AA	CAGTTTCGTCATTTATACAAAATTACATCTATTTTAAAGGAAACAAATATAACG	1643
Db	421	AA	CAGTTTCGTCATTTATACAAAATTACATCTATTTTAAAGGAAACAAATATAACG	480
QY	1644	AT	CCAGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGCTCCTATCAAGAGGATAGC	1703
Db	481	AT	CCAGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGCTCCTATCAAGAGGATAGC	540
QY	1704	AA	TTACAGTGGTGAATGGGAAGTTCTTGGCGCATGATCTTAATATCCAAAGTTGATGA	1763
Db	541	AA	TTACAGTGGTGGTGAATGGGAAGTTCTTGGCGCATGATCTTAATATCCAAAGTTGATGA	600
QY	1764	AG	TCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGTA	1823
Db	601	AG	TCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGTA	660
QY	1824	CG	TAGTACGTTACGTAATTCCTGGAGAGTGACAAAGCTGACCTGGCTACTCACA	1883
Db	661	CG	TAGTACGTTAGTAATTCCTGGAGAGTGACAAAGCTGACCTGGCTACTCACA	720
QY	1884	TT	CTTCTGTCATCAGTCAGCACTGTCCTTTTATAAGTAAGTATAGTAACACAGAGAA	1943
Db	721	TT	CTTCTGTCATCAGTCAGCACTGTCCTTTTATAAGTAAGTATAGTAACACAGAGAA	780
QY	1944	TC	CACACTGTGTCTCCCTTTTACAGCTATCAAGCTCTGGAAGTGAACCAACTGGCAAC	2003
Db	781	TC	CACACTGTGTCTCCCTTTTACAGCTATCAAGCTCTGGAAGTGAACCAACTGGCAAC	840
QY	2004	AA	AGGAATTTTGGGCCACCATTTTGGATTCAGAGGCTCCTCTTCTGACTATCTCTCC	2063
Db	841	AA	AGGAATTTTGGGCCACCATTTTGGATTCAGAGGCTCCTCTTCTGACTATCTCTCC	900
QY	2064	AG	AAATTTTCTTTTTTGAAGTACTACTGGAATTAATGATGGGATGCTCTACAGCC	2123
Db	901	AG	AAATTTTCTTTTTTGAAGTACTACTGGAATTAATGATGGGATGCTCTACAGCC	960
QY	2124	TC	ATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCA	2183
Db	961	TC	ATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCA	1020
QY	2184	GGT	2186	
Db	1021	GGT	1023	

```
QY 2407 CGTGTGGGATCCACGGCTGGTCTCTATGGAGATACCTCTCCCTGATGGCATTAATGCGAG 2466
Db 1877 CGTGTGGGATCCACGGCTGGTCTCTATGGAGATACCTCTCCCTGATGGCATTAATGCGAG 1936
QY 2467 AGGTCAAGATATCTTCAGGGTTGCTATTGCTGGGGCCCACTGCTCTGTGGATCTTCTAT 2526
Db 1937 AGGTCAAGATATCTTCAGGGTTGCTATTGCTGGGGCCCACTGCTCTGTGGATCTTCTAT 1996
QY 2527 GATACAGGATACACGGAACGTTATATGGGTCAACCTGACCAAGATGAAACAGGCTATTAC 2586
Db 1997 GATACAGGATACACGGAACGTTATATGGGTCAACCTGACCAAGATGAAACAGGCTATTAC 2056
QY 2587 TTAGGATCTGTGGCCATGACGACGAGAAAGTTCCCTC 2624
Db 2057 TTAGGATCTGTGGCCATGACGACGAGAAAGTTCCCTC 2094

RESULT 33
AR631282
LOCUS AR631282 1356 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 10 from patent US 6844180.
ACCESSION AR631282
VERSION AR631282.1 GI:59770925
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1356)
AUTHORS Qi,S., Akineanya,K.O., Riviere,P.J.M. and Junien,J.-L.
TITLES Serine protease genes related to DppIV
JOURNAL Patent: US 6844180-A 10 18-JAN-2005;
Ferring BV; ;
NLX;

FEATURES
source Location/Qualifiers
1..1356
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 28.1%; Score 877; DB 2; Length 1356;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 927; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGTGCTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTCTTCTAGTGCCG 60
Db 1 AAGTGCTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTCTTCTAGTGCCG 60
QY 61 CGTTCCGCCCTGGGTTGTCAACGGCGCGCGCGGAGGAGCCACTGCAACAGGACCG 120
Db 61 CGTTCCGCCCTGGGTTGTCAACGGCGCGCGCGGAGGAGCCACTGCAACAGGACCG 120
QY 121 GAGTGGAGCGCGCGAGCATGAGCGCGCGCGGCGGCGGCTCCATAGCGCACGTCGGGACG 180
Db 121 GAGTGGAGCGCGCGAGCATGAGCGCGCGCGGCGGCGGCTCCATAGCGCACGTCGGGACG 180
QY 181 TCGGGCGGGCGCGGGGAGGAAATGCAACATGCGACGACCAATGAAACAGAACAG 240
Db 181 TCGGGCGGGCGCGGGGAGGAAATGCAACATGCGACGACCAATGAAACAGAACAG 240
QY 241 CTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTAATCAACAGGATCGG 300
Db 241 CTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTAATCAACAGGATCGG 300
QY 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCT 360
Db 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCT 360
QY 361 GCCGATACAGAAATATCATGGCTACATGATGGCTTAAGGCCACCATGATTTTCATGTT 420
Db 361 GCCGATACAGAAATATCATGGCTACATGATGGCTTAAGGCCACCATGATTTTCATGTT 420
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTATTACCTTGCCATGCT 480
```

```
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTATTACCTTGCCATGCT 480
QY 481 GGTGAGAAACAGAGAAATACACTGTTTATTTCTGAATTCCTGAATTCCTGAATTCCTGA 540
Db 481 GGTGAGAAACAGAGAAATACACTGTTTATTTCTGAATTCCTGAATTCCTGAATTCCTGA 540
QY 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
QY 601 GGAATGTATCTCCGAGAGAGAACTATTAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
Db 601 GGAATGTATCTCCGAGAGAGAACTATTAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
QY 661 ATTGCTCTTACGATATATACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGCTAGTGA 720
Db 661 ATTGCTCTTACGATATATACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGCTAGTGA 720
QY 721 ATTATACCTTAAGATGAGGGCCACAGGATTTACGCAACACTTTAAGGCCCAAT 780
Db 721 ATTATACCTTAAGATGAGGGCCACAGGATTTACGCAACACTTTAAGGCCCAAT 780
QY 781 CTAGTGGAACTAGTCTCCCAACATACGATCGATCCAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGGAACTAGTCTCCCAACATACGATCGATCCAAATTTATGCCCGCTGATCCA 840
QY 841 GACTGGATTTGTTTATACATAGCAACGATATTGGATATCTAAACATCGTAACCCAGAGAA 900
Db 841 GACTGGATTTGTTTATACATAGCAACGATATTGGATATCTAAACATCGTAACCCAGAGAA 900
QY 901 GAAAGGAGACTCACTTATGTGCACAATG 928
Db 901 GAAAGGAGACTCACTTATGTGCACAATG 928

RESULT 34
AX608733
LOCUS AX608733 1356 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 10 from Patent WO0231134.
ACCESSION AX608733
VERSION AX608733.1 GI:28404302
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1
AUTHORS Qi,S., Akineanya,K.O., Riviere,P.J. and Junien,J.L.
TITLES Novel serine proteases, genes related to dppiv
JOURNAL Patent: WO 0231134-A 10 18-APR-2002;
Perring BV (NL)

FEATURES
source Location/Qualifiers
1..1356
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 28.1%; Score 877; DB 2; Length 1356;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 927; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGTGCTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTTCTAGTGCCG 60
Db 1 AAGTGCTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTTCTAGTGCCG 60
QY 61 CGTTCCGCCCTGGGTTGTCAACGGCGCGCGGAGGAGCCACTGCAACAGGACCG 120
Db 61 CGTTCCGCCCTGGGTTGTCAACGGCGCGCGGAGGAGCCACTGCAACAGGACCG 120
QY 121 GAGTGGAGCGCGCGAGCATGAGCGCGCGCGGCGGCGGCTCCATAGCGCACGTCGGGACG 180
Db 121 GAGTGGAGCGCGCGAGCATGAGCGCGCGCGGCGGCGGCTCCATAGCGCACGTCGGGACG 180
QY 181 TCGGGCGGGCGCGGGGAGGAAATGCAACATGCGACGACCAATGAAACAGAACAG 240
Db 181 TCGGGCGGGCGCGGGGAGGAAATGCAACATGCGACGACCAATGAAACAGAACAG 240
QY 241 CTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTAATCAACAGGATCGG 300
Db 241 CTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTAATCAACAGGATCGG 300
QY 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCT 360
Db 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCT 360
QY 361 GCCGATACAGAAATATCATGGCTACATGATGGCTTAAGGCCACCATGATTTTCATGTT 420
Db 361 GCCGATACAGAAATATCATGGCTACATGATGGCTTAAGGCCACCATGATTTTCATGTT 420
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTATTACCTTGCCATGCT 480
```


Thu Jun 22 09:04:25 2006

FEATURES
source

Location/Qualifiers
1..1083
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q22"
/clone="scPBMCd3-3-1"
/cell_type="PHA stimulated PBMC"
<1..>1083
/gene="DPP8"
<1..>1083
/gene="DPP8"
/note="serine protease; member of prolyl oligopeptidase S9b family belonging to enzyme clan SC; additional splice variants of this mRNA are found in GenBank Accession Numbers AF221634, AF221635 and AF221636"
/codon_start=2
/product="dipeptidyl peptidase 8"
/protein_id="AAG29769.1"
/translation="EEDARSAGVATFVLQEFDRYSGYWHCPKAETTPSGKILRLYL
ENDSEVEIIVHTSPMLSTRADSFYKPTGTANPKVTFKMSIMIDAGRLIIVDEV
RLIVFEGTKDPSLEHLVYVYNPGEVTLTRDGYSHSCCISQHCDFIFSKYSNOK
NPHCVSLYKLSPEDDPTKTEFWATILDSAGLPDYTPPEIFSFSTFTGLYGLM
YKPHDLQPKKPTVLFYGGPOVLNNRPFKVKYFRLATLSLGVVVVVVINDRSC
HRGLKPEGAFKYMGEIIDDQVGLQYLSRYDFIDLDVRVGHWSYGGVYLSLMALM
QRSIDIPRVAIAGAPVTLWIFDYDTGT"

gene

CDS

415 CTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAA 474
1935 CCAGAAGAACCCACACTGTGTGCTCCCTTTACAAAGCTATCAAGTCTCTGAAGATGACCCCAAC 1994
475 CCAGAAGAACCCACACTGTGTGCTCCCTTTACAAAGCTATCAAGTCTCTGAAGATGACCCCAAC 534
1995 TTGCAAAACAAAGAAATTTTGGGCGCCACATTTTGGATTCAGCAGGTCTCTTCTCTGACTA 2054
535 TTGCAAAACAAAGAAATTTTGGGCGCCACATTTTGGATTCAGCAGGTCTCTTCTCTGACTA 594
2055 TACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTACATTTTACATTTGATGGATGCT 2114
595 TACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTACATTTTACATTTGATGGATGCT 654
2115 CTACAGGCTCATGATCTACAGCTCGAAGAAATATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2174
655 CTACAGGCTCATGATCTACAGCTCGAAGAAATATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 714
2175 TGGTCTCTCAGGTGCAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAA 2234
715 TGGTCTCAGGTGCAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAA 774
2235 TACCTTAGCTCTCTAGTTATGCTGTAGTGTAGTATAGTATAGTATAGTATAGTATAGTATAGT 2294
775 TACCTTAGCTCTCTAGTTATGCTGTAGTGTAGTATAGTATAGTATAGTATAGTATAGTATAGT 834
2295 AGGGCTTAAATTTGAAGCGCTTTAAATATAAATGGGTCAATAGAAATTCAGCATCA 2354
835 AGGGCTTAAATTTGAAGCGCTTTAAATATAAATGGGTCAATAGAAATTCAGCATCA 894
2355 GGTGGAAGGACTCCATATCTAGTCTTCGATATGATTTCAATTCATTCATTCATTCATTCATTC 2414
895 GGTGGAAGGACTCCATATCTAGTCTTCGATATGATTTCAATTCATTCATTCATTCATTCATTC 954
2415 CATCCAGGCTGTCTATGAGGATACCTCTCCCTGATGTCATTAATTCAGAGGTGAGT 2474
955 CATCCAGGCTGTCTATGAGGATACCTCTCCCTGATGTCATTAATTCAGAGGTGAGT 1014
2475 TATCTTCAGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGG 2534
1015 TATCTTCAGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGG 1074
2535 ATACACGGA 2543
1075 ATACACGGA 1083

RESULT 38
AF221637 1083 bp mRNA linear PRI 05-NOV-2000
LOCUS Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, partial cds,
DEFINITION alternatively spliced.
ACCESSION AF221637
VERSION AF221637.1 GI:11095193
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1083)
AUTHORS Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W.
and Gorrell,M.D.
TITLE Cloning, expression and chromosomal localization of a novel human
dipeptidyl peptidase (DPP) IV homolog, DPP8
JOURNAL Eur. J. Biochem. 267 (20), 6140-6150 (2000)
PUBMED 11012666
REFERENCE 2 (bases 1 to 1083)
AUTHORS Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver
Centre, Centenary Institute of Cell Biology and Cancer Medicine,
Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia

ORIGIN

Query Match 25.3%; Score 789; DB 5; Length 1083;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1755 AGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGCACCAAGACTCCCTTTAGAGCA 1814
DB 295 AGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGCACCAAGACTCCCTTTAGAGCA 354
QY 1815 TCACCTGTACGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGCTGACCTGACCGTGG 1874
DB 355 TCACCTGTACGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGCTGACCTGACCGTGG 414
QY 1875 CTATCACAATCTTGTGTCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAA 1934
DB 415 CTACTCACAATCTTGTGTCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAA 474
QY 1935 CCAGAAGAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCAAC 1994
DB 475 CCAGAAGAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCAAC 534
QY 1995 TTGCAAAAACAAAGAAATTTTGGGCGCCACATTTTGGATTCAGCAGGTCTCTTCTCTGACTA 2054
DB 535 TTGCAAAAACAAAGAAATTTTGGGCGCCACATTTTGGATTCAGCAGGTCTCTTCTCTGACTA 594
QY 2055 TACTCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACATTTGATGGATGCT 2114
DB 595 TACTCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACATTTGATGGATGCT 654
QY 2115 CTACAGGCTCATGATCTACAGCTCGAAGAAATATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2174
DB 655 CTACAGGCTCATGATCTACAGCTCGAAGAAATATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 714
QY 2175 TGGTCTCTCAGGTGCAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAA 2234
DB 715 TGGTCTCTCAGGTGCAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAA 774
QY 2235 TACCTTAGCTCTCTAGTTATGCTGTAGTGTAGTATAGTATAGTATAGTATAGTATAGTATAGT 2294
DB 775 TACCTTAGCTCTCTAGTTATGCTGTAGTGTAGTATAGTATAGTATAGTATAGTATAGTATAGT 834
QY 2295 AGGGCTTAAATTTGAAGCGCTTTAAATATAAATGGGTCAATAGAAATTCAGCATCA 2354
DB 835 AGGGCTTAAATTTGAAGCGCTTTAAATATAAATGGGTCAATAGAAATTCAGCATCA 894

```
Qy 2355 GGTGGAAGACCTCAATATCTAGCTTCTCGATATGATTTCAATGACTTATGATCGTGTGG 2414
Db 895 GGTGGAAGACCTCAATATCTAGCTTCTCGATATGATTTCAATGACTTATGATCGTGTGG 954
Qy 2415 CATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGCAATTAATCGAGGCTCAGA 2474
Db 955 CATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGCAATTAATCGAGGCTCAGA 1014
Qy 2475 TATCTTCAGGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGATCTTCTATGATACAG 2534
Db 1015 TATCTTCAGGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGATCTTCTATGATACAG 2534
Qy 2535 ATACACGGA 2543
Db 1075 ATACACGGA 1083

RESULT 39
AR631286
LOCUS AR631286 832 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 18 from patent US 644180.
ACCESSION AR631286
VERSION AR631286.1 GI:59770932
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 832)
AUTHORS O.I.S., Akineanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 644180-A 18 JAN-2005;
Ferring BV;
NLX;
FEATURES
source Location/Qualifiers
1..832
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 24.4%; Score 760; DB 2; Length 832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTACTGCGCGCTGCTTCTTAGTGC 60
Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTACTGCGCGCTGCTTCTTAGTGC 60
Qy 61 CGTTCCGCCCTCGGCTTGTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CGTTCCGCCCTCGGCTTGTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Qy 121 GAGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 GAGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Qy 181 TCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 TCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Qy 241 CTGGGTGTTGAGATTTGAACTGCGGACTGTGAGGAGAAATGCAATGCGAGCAATGGAATCAGGATCG 300
Db 241 CTGGGTGTTGAGATTTGAACTGCGGACTGTGAGGAGAAATGCAATGCGAGCAATGGAATCAGGATCG 300
Qy 301 CCTAAATTTGAGCCCTTTTATGTTGAGCGGTATTTCTTGGAGTCAAGTTTAAAGTGTCTT 360
Db 301 CCTAAATTTGAGCCCTTTTATGTTGAGCGGTATTTCTTGGAGTCAAGTTTAAAGTGTCTT 360
Qy 361 GCCGATACCGAATAATCATGCTACATGATGCTTAAGGCCACCATGATTTTATGTTT 420
Db 361 GCCGATACCGAATAATCATGCTACATGATGCTTAAGGCCACCATGATTTTATGTTT 420
Qy 421 GTGAAGAGGATGATGCAGATGACCTCATTCAGACAGAAATCTATTACCTTGGCATGTCT 480
```

```
Db 421 GTGAAGAGGATGATGCAGATGACCTCATTCAGACAGAAATCTATTACCTTGGCATGTCT 480
Qy 481 GGTGAGAAACAGAGAAAATACACTGTTTATTCTGAAATTCCTCAAAATCATCATAGAGCA 540
Db 481 GGTGAGAAACAGAGAAAATACACTGTTTATTCTGAAATTCCTCAAAATCATCATAGAGCA 540
Qy 541 GCAGTCTTAAATGCTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAAATGCTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGTATTCTCGAGAGAGAACTATTAAAGAGAAAACGCAATGGAACAGTCGGA 660
Db 601 GGAATGTATTCTCGAGAGAGAACTATTAAAGAGAAAACGCAATGGAACAGTCGGA 660
Qy 661 ATTGCTTCTTACGATTATCACAAGGAGTGGAAACATTTCTGTTTCAAGCCGTTAGTGA 720
Db 661 ATTGCTTCTTACGATTATCACAAGGAGTGGAAACATTTCTGTTTCAAGCCGTTAGTGA 720
Qy 721 ATTATCATGTAAGATGAGAGGCCACAGGATTTACGC 760
Db 721 ATTATCATGTAAGATGAGAGGCCACAGGATTTACGC 760

RESULT 40
AR608741
LOCUS AR608741 832 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 18 from Patent WO0231134.
ACCESSION AR608741
VERSION AR608741.1 GI:28404306
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Qi, S., Akineanya, K.O., Riviere, P.J. and Junien, J.L.
TITLE Novel serine protease genes related to DPPIV
JOURNAL Patent: WO 0231134-A 18 APR-2002;
Ferring BV (NL)
FEATURES
source Location/Qualifiers
1..832
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 24.4%; Score 760; DB 2; Length 832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGC 60
Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGC 60
Qy 61 CGTTCCGCCCTCGGCTTGTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CGTTCCGCCCTCGGCTTGTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Qy 121 GAGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 GAGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Qy 181 TCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 TCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Qy 241 CTGGGTGTTGAGATTTGAACTGCGGACTGTGAGGAGAAATGCAATGCGAGCAATGGAATCAGGATCG 300
Db 241 CTGGGTGTTGAGATTTGAACTGCGGACTGTGAGGAGAAATGCAATGCGAGCAATGGAATCAGGATCG 300
Qy 301 CCTAAATTTGAGCCCTTTTATGTTGAGCGGTATTTCTTGGAGTCAAGTTTAAAGTGTCTT 360
```

301 CCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360
361 GCCGATACAGAAAATATCATGCTACATGCTAGCTAGGCTAAGCGCACCATGATTTTCATGTTT 420
361 GCCGATACAGAAAATATCATGCTACATGCTAGGCTAAGCGCACCATGATTTTCATGTTT 420
421 GTGAGAGGAGTATGATCCAGATGGACCTCATTCAGACAGATCATTTACCTTGCCATGCT 480
421 GTGAGAGGAGTATGATCCAGATGGACCTCATTCAGACAGATCATTTACCTTGCCATGCT 480
481 GGTGAGAACAGAGAAAATACACTGTTTATTCGAAATTCCTCCAAACTATCAATAGAGCA 540
481 GGTGAGAACAGAGAAAATACACTGTTTATTCGAAATTCCTCCAAACTATCAATAGAGCA 540
541 CGAGTCTTAATGCTCTCTTGAAGCCTCTTTGGATCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
541 CGAGTCTTAATGCTCTCTTGAAGCCTCTTTGGATCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
601 GGAATGATTTCTCGAAGAGAACTATTAAGAGAAAAGAAAAGCAATGGAACTGGACGCA 660
601 GGAATGATTTCTCGAAGAGAACTATTAAGAGAAAAGAAAAGCAATGGAACTGGACGCA 660
661 ATTGCTTCTTACGATATCACCAAGAGTGGACATTTCTTTCAAGCCGGTAGTGGA 720
661 ATTGCTTCTTACGATATCACCAAGAGTGGACATTTCTTTCAAGCCGGTAGTGGA 720
721 ATTTATCAGTAAAGATGGAGGCGCACAGGATTTACGC 760
721 ATTTATCAGTAAAGATGGAGGCGCACAGGATTTACGC 760

RESULT 41
AX083130
LOCUS AX083130 873 bp DNA linear PAT 28-FEB-2001
DEFINITION Sequence 33 from Patent WO0110903.
ACCESSION AX083130
VERSION AX083130.1 GI:13185037
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
Yue,H., Lal,P., Tang,Y.T., Bandman,O., Baughn,M.R., Azimzai,Y.,
Lu,D.A. and Yang,J.
Proteases and protease inhibitors
Patent: WO 0110903-A 33 15-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1. .873
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 1903112CB1"

ORIGIN
Query Match 23.7%; Score 739; DB 2; Length 873;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 789; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2331 GGGTCAAAATAGAAATGACGATCAGGTGGAGGACCTCCATATCTAGCTTCTCGATATGA 2390
83 GGGTCAAAATAGAAATGACGATCAGGTGGAGGACCTCCATATCTAGCTTCTCGATATGA 142
2391 TTTTCATTCGATTTAGATCTGTGGGATCCACGGCTGCTCTATGGAGGATACCTTCCT 2450
143 TTTTCATTCGATTTAGATCTGTGGGATCCACGGCTGCTCTATGGAGGATACCTTCCT 202
2451 GATGGCATTAATGACAGAGGTGAGATATCTTCAGGGTGTCTATTCCTGGGCCCCAGTAC 2510
203 GATGGCATTAATGACAGAGGTGAGATATCTTCAGGGTGTCTATTCCTGGGCCCCAGTAC 262

2511 TCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTCAACCTGACAGAA 2570
263 TCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTCAACCTGACAGAA 322
2571 TGAACAGAGGCTATTATCTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTCAAC 2630
323 TGAACAGAGGCTATTATCTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTCAAC 382
2631 AAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGTCCATTTTGCACATACAG 2690
383 AAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGTCCATTTTGCACATACAG 442
2691 TATATTACTGAGCTTTTATGAGGCTGGAAAGCCATGATTTTACAGATCTATCTCA 2750
443 TATATTACTGAGCTTTTATGAGGCTGGAAAGCCATGATTTTACAGATCTATCTCA 502
2751 GGAGAGACACAGATTAAGATTTCTGATCGGAGAACATTAATGATCTTTTGGCA 2810
503 GGAGAGACACAGATTAAGATTTCTGATCGGAGAACATTAATGATCTTTTGGCA 562
2811 CTACCTTCAAGAAAACCTTGGATCACGTTATGCTGCTCTAAAAGTGTATTAATTTGACC 2870
563 CTACCTTCAAGAAAACCTTGGATCACGTTATGCTGCTCTAAAAGTGTATTAATTTGACC 622
2871 TGTGTAGAACTCTCTGTTATACATGCTGGCTATTTAAACCAATGAGGAGTTTAAATCAAC 2930
623 TGTGTAGAACTCTCTGTTATACATGCTGGCTATTTAAACCAATGAGGAGTTTAAATCAAC 682
2931 AAAACACAGAAATGATCATCATTTTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2990
683 AAAACACAGAAATGATCATCATTTTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
2991 AAATGTGCTGCCATGAGGCTCTAGGTTTGTGTTAGTAAATCTAATACCTTAACCCAC 3050
743 AAATGTGCTGCCATGAGGCTCTAGGTTTGTGTTAGTAAATCTAATACCTTAACCCAC 802
3051 ATGCTCAAAATCAAAATGATCATTTCTGAGAGCCCGAGCAATACCATTAAGATTAATA 3110
803 ATGCTCAAAATCAAAATGATCATTTCTGAGAGCCCGAGCAATACCATTAAGATTAATA 862
3111 AAAAAGAAAA 3120
863 AAAAAGAAAA 872

RESULT 42
AX524942
LOCUS AX524942 735 bp DNA linear PAT 21-NOV-2002
DEFINITION Sequence 15 from Patent WO02066627.
ACCESSION AX524942
VERSION AX524942.1 GI:25170024
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
Liou,J.
Regulation of human dipeptidyl peptidase 8
Patent: WO 02066627-A 15 29-AUG-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source
1. .735
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 22.7%; Score 708; DB 2; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2274	CAACAGGGATCCTGTCAACCGAGGGCTTAAATTTGAAGGGCCCTTTAAATATATAAATGGG	2333
DB	1	CAACAGGGGATCCTGTCAACCGAGGGCTTAAATTTGAAGGGCCCTTTAAATATATAAATGGG	60
QY	2334	TCAAATAGAAAAATTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTT	2393
DB	61	TCAAATAGAAAAATTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTT	120
QY	2394	CATTGACTTTAGATCTGTGGGCACTCCACGGCTGGCTCATGAGGATACCTCTCCCTGAT	2453
DB	121	CATTGACTTTAGATCTGTGGGCACTCCACGGCTGGCTCATGAGGATACCTCTCCCTGAT	180
QY	2454	GCATTAAATGACAGGTTCAGATATCTTCAGGGTTCCTATGCTGGGGCCCCCAGTCACTCT	2513
DB	181	GCATTAAATGACAGGTTCAGATATCTTCAGGGTTCCTATGCTGGGGCCCCCAGTCACTCT	240
QY	2514	GTGATCTTCTATGATACAGATACACGGAACTGTTATATGGGTCACCTGCACAGAAATGA	2573
DB	241	GTGATCTTCTATGATACAGATACACGGAACTGTTATATGGGTCACCTGCACAGAAATGA	300
QY	2574	ACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAA	2633
DB	301	ACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAA	360
QY	2634	TCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGTCCTATTTGCACATACCAATAT	2693
DB	361	TCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGTCCTATTTGCACATACCAATAT	420
QY	2694	ATTACTGAGTTTTTACTGAGGGCTGGAAGCCCATATGATTTACAGATCTATCTCAGGA	2753
DB	421	ATTACTGAGTTTTTACTGAGGGCTGGAAGCCCATATGATTTACAGATCTATCTCAGGA	480
QY	2754	GAGACAGCATAGAGTTCTCTGAATCGGAGAAACATTATGAACTGCACTTTTGGCACTA	2813
DB	481	GAGACAGCATAGAGTTCTCTGAATCGGAGAAACATTATGAACTGCACTTTTGGCACTA	540
QY	2814	CCTTTCAAGAAAACTTTGATCAGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2873
DB	541	CCTTTCAAGAAAACTTTGATCAGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	600
QY	2874	GTGAACCTCTCTGTATACACTGGCTATTACCAAAATGAGAGGTTTAAATCAACAGAA	2933
DB	601	GTGAACCTCTCTGTATACACTGGCTATTACCAAAATGAGAGGTTTAAATCAACAGAA	660
QY	2934	ACACAGAAATGATCATCATATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2981
DB	661	ACACAGAAATGATCATCATATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	708
RESULT 43			
AX524935			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			


```
RESULT 46
AX869790
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Ota, T., Iwagai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, F., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primer for synthesising full-length cDNA and their use
Patent: EP 1074617-A 4695 07-FEB-2001;
Research Association for Biotechnology (JP)
Location/Qualifiers
1..631
/organism="Homo sapiens"
/db_xref="taxon:9606"
ORIGIN
Query Match
Best Local Similarity 16.1%; Score 501; DB 2; Length 631;
Matches 551; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 832 GCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATTGGATATCTAACATCGTA 891
Db 37 GCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATTGGATATCTAACATCGTA 96
QY 892 ACCAGAGAAAGAGAGACTCACTTATGTGCATAGTAGTAGCCACATGGAGAGAT 951
Db 97 ACCAGAGAAAGAGAGACTCACTTATGTGCATAGTAGTAGCCACATGGAGAGAT 156
QY 952 GCCAGATCAGCTGGAGTCCCTACCTTTGTTCTCCAGAGAGATTTGTAGATATTCGGC 1011
Db 157 GCCAGATCAGCTGGAGTCCCTACCTTTGTTCTCCAGAGAGATTTGTAGATATTCGGC 216
QY 1012 TATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGAAATCTTA 1071
Db 217 TATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGAAATCTTA 276
QY 1072 TATGAAGAAATGATGAAATCTGAGGTGGAATTTATGTTATCATCCCTATGTTGGAA 1131
Db 277 TATGAAGAAATGATGAAATCTGAGGTGGAATTTATGTTATCATCCCTATGTTGGAA 336
QY 1132 ACAAGGAGGCGAGATTCATTCGGTTATCCTTAAACAGGTACAGAAATCTTAAAGTCACT 1191
Db 337 ACAAGGAGGCGAGATTCATTCGGTTATCCTTAAACAGGTACAGAAATCTTAAAGTCACT 396
QY 1192 TTTAGATGTGCAAAATATGATTTGATGCTGAAGGAGGATCATAGATGTCTAGATAAG 1251
Db 397 TTTAGATGTGCAAAATATGATTTGATGCTGAAGGAGGATCATAGATGTCTAGATAAG 456
QY 1252 GAACCTAATCAACCTTTTGAGATTTCTATTTGAAGGAGTGAATATATTGCCAGAGCTGA 1311
Db 457 GAACCTAATCAACCTTTTGAGATTTCTATTTGAAGGAGTGAATATATTGCCAGAGCTGA 516
QY 1312 TGGACTCTGAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCCTA 1371
Db 517 TGGACTCTGAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCCTA 576
QY 1372 CAGATAGTGTG 1383
Db 577 CAGATAGTGTG 588
```

RESULT 47

```
AR631285
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
1 (bases 1 to 620)
AUTHORS
O. S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE
Serine protease genes related to DPPIV
JOURNAL
Patent: US 6844180-A 16 18-JAN-2005;
Ferring BV;;
NLX,
Location/Qualifiers
1..620
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match
Best Local Similarity 15.1%; Score 472; DB 2; Length 620;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGTCTTAAGCCTCCGAGCCCAAGGCCGCTGCTACTGCGCGCTGCTTTCTTAGTGGCG 60
Db 1 AAGTCTTAAGCCTCCGAGCCCAAGGCCGCTGCTACTGCGCGCTGCTTTCTTAGTGGCG 60
QY 61 CGTTGCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CGTTGCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 GAGTGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 GAGTGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 181 TCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 TCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGACTGCTGAGGAGATATTTGAATCATAGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGACTGCTGAGGAGATATTTGAATCATAGATCGG 300
QY 301 CCTAAATTGAGGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGCTGCTT 360
Db 301 CCTAAATTGAGGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGCTGCTT 360
QY 361 GCCGATACCAGAAATATCATGCTGCTACATGATGGCTAAGGCCACCATGATTTTATGTTT 420
Db 361 GCCGATACCAGAAATATCATGCTGCTACATGATGGCTAAGGCCACCATGATTTTATGTTT 420
QY 421 GTCAAGAGGAATGATGAGATGAGCTCATTTTCAGACAGAAATCTATTACCTTG 472
Db 421 GTCAAGAGGAATGATGAGATGAGCTCATTTTCAGACAGAAATCTATTACCTTG 472
RESULT 48
AX608739
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
Novel serine protease genes related to dppiv
ORIGIN
Query Match
Best Local Similarity 15.1%; Score 472; DB 2; Length 620;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGTCTTAAGCCTCCGAGCCCAAGGCCGCTGCTACTGCGCGCTGCTTTCTTAGTGGCG 60
Db 1 AAGTCTTAAGCCTCCGAGCCCAAGGCCGCTGCTACTGCGCGCTGCTTTCTTAGTGGCG 60
QY 61 CGTTGCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CGTTGCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 GAGTGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 GAGTGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 181 TCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 TCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGACTGCTGAGGAGATATTTGAATCATAGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGACTGCTGAGGAGATATTTGAATCATAGATCGG 300
QY 301 CCTAAATTGAGGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGCTGCTT 360
Db 301 CCTAAATTGAGGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGCTGCTT 360
QY 361 GCCGATACCAGAAATATCATGCTGCTACATGATGGCTAAGGCCACCATGATTTTATGTTT 420
Db 361 GCCGATACCAGAAATATCATGCTGCTACATGATGGCTAAGGCCACCATGATTTTATGTTT 420
QY 421 GTCAAGAGGAATGATGAGATGAGCTCATTTTCAGACAGAAATCTATTACCTTG 472
Db 421 GTCAAGAGGAATGATGAGATGAGCTCATTTTCAGACAGAAATCTATTACCTTG 472
RESULT 49
AX608739
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
Novel serine protease genes related to dppiv
ORIGIN
Query Match
Best Local Similarity 15.1%; Score 472; DB 2; Length 620;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGTCTTAAGCCTCCGAGCCCAAGGCCGCTGCTACTGCGCGCTGCTTTCTTAGTGGCG 60
Db 1 AAGTCTTAAGCCTCCGAGCCCAAGGCCGCTGCTACTGCGCGCTGCTTTCTTAGTGGCG 60
QY 61 CGTTGCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CGTTGCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 GAGTGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 GAGTGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 181 TCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 TCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGACTGCTGAGGAGATATTTGAATCATAGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGACTGCTGAGGAGATATTTGAATCATAGATCGG 300
QY 301 CCTAAATTGAGGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGCTGCTT 360
Db 301 CCTAAATTGAGGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGCTGCTT 360
QY 361 GCCGATACCAGAAATATCATGCTGCTACATGATGGCTAAGGCCACCATGATTTTATGTTT 420
Db 361 GCCGATACCAGAAATATCATGCTGCTACATGATGGCTAAGGCCACCATGATTTTATGTTT 420
QY 421 GTCAAGAGGAATGATGAGATGAGCTCATTTTCAGACAGAAATCTATTACCTTG 472
Db 421 GTCAAGAGGAATGATGAGATGAGCTCATTTTCAGACAGAAATCTATTACCTTG 472
```

JOURNAL	Patent: WO 0231134-A 16 18-APR-2002;
FEATURES	Ferring BV (NL) Location/Qualifiers 1. .620 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
ORIGIN	Query Match 15.1%; Score 472; DB 2; Length 620; Best Local Similarity 100.0%; Pred. No. 1.4e-265; Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 AAGTCTAAAGCCCTCCGAGGCCAGCGCGTGTACTTGCAGCGCGTCTTCTTAGTCGCG 60
Db	1 AAGTCTAAAGCCCTCCGAGGCCAGCGCGTGTACTTGCAGCGCGTCTTCTTAGTCGCG 60
QY	61 CGTTCGCCGCTCGGTGTGTCAACCGCGCCGCGCGAGGAGCCACTGCAACGAGGACCG 120
Db	61 CGTTCGCCGCTCGGTGTGTCAACCGCGCCGCGCGAGGAGCCACTGCAACGAGGACCG 120
QY	121 GAGTGGAGCGCGCGCAGCATGAAGCGCGAGCGCGCGCTCCATAGCGCAGCTCGGGACGG 180
Db	121 GAGTGGAGCGCGCGCAGCATGAAGCGCGAGCGCGCGCTCCATAGCGCAGCTCGGGACGG 180
QY	181 TCCGGCGCGCGCGCGGGAAGGAATGCAACATGCGCAGCAATGGAACAGAACACAG 240
Db	181 TCCGGCGCGCGCGCGGGAAGGAATGCAACATGCGCAGCAATGGAACAGAACACAG 240
QY	241 CTGGGTGTGAGATATTTGAAACTCGGACCTGTGAGGAGAAATTTGAATCAGAGATCGG 300
Db	241 CTGGGTGTGAGATATTTGAAACTCGGACCTGTGAGGAGAAATTTGAATCAGAGATCGG 300
QY	301 CCTAAATTCGAGCCCTTTTATGTTGAGCGGTATTCCTCGAGTCAGCTTAAAGCTGCTT 360
Db	301 CCTAAATTCGAGCCCTTTTATGTTGAGCGGTATTCCTCGAGTCAGCTTAAAGCTGCTT 360
QY	361 GCGATACAGAAATATCATGCTACATGATGCTAGGACCAACACATGATTTTCATGTTT 420
Db	361 GCGATACAGAAATATCATGCTACATGATGCTAGGACCAACACATGATTTTCATGTTT 420
QY	421 GTGAAGGAGGATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTG 472
Db	421 GTGAAGGAGGATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTG 472
RESULT 49	
AC105129	198295 bp DNA linear PRI 26-APR-2002
LOCUS	Homo sapiens chromosome 15, clone RP11-349G13, complete sequence.
DEFINITION	AC105129
ACCESSION	AC105129.4 GI:20330981
VERSION	HTG.
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 198295)
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE	Homo sapiens chromosome 15, clone RP11-349G13
JOURNAL	Unpublished
AUTHORS	2 (bases 1 to 198295)
REFERENCE	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczkzy,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.


```
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/19, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1..587
FT /organism="Homo sapiens (human)".
FEATURES
source
Location/Qualifiers
1..587
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 6.6%; Score 205; DB 2; Length 587;
Best Local Similarity 99.2%; Pred. No. 5.1e-108;
Matches 355; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2655 TTTCCTGGATGAGAAATGTCATTTTGCACATACCATGATATTTAGTGGATGAGTTC 2714
DB 455 TTTCCTGGATGAGAAATGTCATTTTGCACATACCATGATATTTAGTGGATGAG 396
QY 2715 GGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCC 2774
DB 395 GGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCC 336
QY 2775 TGAATCGGAGAACATTTATGAACTGTCATCTTTTGGACCTGTAGAACTCTCTGGTATACAC 2834
DB 335 TGAATCGGAGAACATTTATGAACTGTCATCTTTTGGACCTGTAGAACTCTCTGGATC 276
QY 2835 ACGTATTGCTGCTCTAAAGTGATATAATTTTGGACCTGTAGAACTCTCTGGTATACAC 2894
DB 275 ACGTATTGCTGCTCTAAAGTGATATAATTTTGGACCTGTAGAACTCTTTTGGTATACAC 216
QY 2895 TGGCTATTTAACCAATGAGAGGTTTAAATCAACAGAAAAACACAGAAATTGATCATCAT 2954
DB 215 TGGCTATTTAACCAATGAGAGGTTTAAATCAACAGAAAAACACAGAAATTGATCATCAT 156
QY 2955 TTTGATACCTGCGCATGTAACATCTACTCTCTGAAAATAAATGTGGTCCATGCAGGGT 3012
DB 155 TTTGATACCTGCGCATGTAACATCTACTCTCTGAAAATAAATGTGGTCCATGCAGGGT 98
RESULT 54
AX970984 308 bp DNA linear PAT 15-JAN-2004
LOCUS
DEFINITION Sequence 1787 from Patent EPI104808.
ACCESSION AX970984
VERSION AX970984.1 GI:40978887
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1
AUTHORS Dumas Milne Edwards, J.B., Jobert, S. and Giordano, J.Y.
TITLE ESTs and encoded human proteins
JOURNAL Patent: EP 1104808-A 1787 06-JUN-2001;
Genset (FR)
FEATURES
source
1..308
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
17..>178
/note="unnamed protein product"
CDS
/codon_start=1
/protein_id="CAFI6092.1"
/db_xref="GI:40978888"
/translation="MQAEKPPSEPNRLLLHGLDENVHPAHTSILLSFLVRAGKHI
VRSILRRDTA"
sig_peptide 17..133
/note="Von Heijne matrix score 4.599999990463257 seq
AHTSILLSFLVRA/GK"
ORIGIN
Query Match 6.0%; Score 188; DB 2; Length 308;
Best Local Similarity 99.7%; Pred. No. 5.4e-98;
Matches 308; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2586 CTTAGATCTGTGGCCATGCAAGCGAAAAAGTTCCTCTGAAACCAATGTTTACTGCT 2645
DB 1 CTTAGATCTGTGGCCATGCAAGCGAAAAAGTTCCTCTGAAACCAATGTTTACTGCT 60
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/19, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1..587
FT /organism="Homo sapiens (human)".
FEATURES
source
Location/Qualifiers
1..587
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 6.6%; Score 205; DB 2; Length 587;
Best Local Similarity 99.2%; Pred. No. 5.1e-108;
Matches 355; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2655 TTTCCTGGATGAGAAATGTCATTTTGCACATACCATGATATTTAGTGGATGAG 2714
DB 455 TTTCCTGGATGAGAAATGTCATTTTGCACATACCATGATATTTAGTGGATGAG 396
QY 2715 GGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCC 2774
DB 395 GGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCC 336
QY 2775 TGAATCGGAGAACATTTATGAACTGTCATCTTTTGGACCTGTAGAACTCTCTGGTATACAC 2834
DB 335 TGAATCGGAGAACATTTATGAACTGTCATCTTTTGGACCTGTAGAACTCTCTGGATC 276
QY 2835 ACGTATTGCTGCTCTAAAGTGATATAATTTTGGACCTGTAGAACTCTCTGGTATACAC 2894
DB 275 ACGTATTGCTGCTCTAAAGTGATATAATTTTGGACCTGTAGAACTCTTTTGGTATACAC 216
QY 2895 TGGCTATTTAACCAATGAGAGGTTTAAATCAACAGAAAAACACAGAAATTGATCATCAT 2954
DB 215 TGGCTATTTAACCAATGAGAGGTTTAAATCAACAGAAAAACACAGAAATTGATCATCAT 156
QY 2955 TTTGATACCTGCGCATGTAACATCTACTCTCTGAAAATAAATGTGGTCCATGCAGGGT 3012
DB 155 TTTGATACCTGCGCATGTAACATCTACTCTCTGAAAATAAATGTGGTCCATGCAGGGT 98
RESULT 53
AX874760/c
AX874760 587 bp DNA linear PAT 17-DEC-2003
LOCUS
DEFINITION Sequence 9665 from Patent EPI074617.
ACCESSION AX874760
VERSION AX874760.1 GI:40029519
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 9665 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source
1..587
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

QY 2646 CTTACATGGTTCTCGGATGAGATGTCATTTTGACATPACCAAGTATATTACTGAGTTT 2705
Db |||||
61 CTTACATGGTTCTCGGATGAGATGTCATTTTGACATPACCAAGTATATTACTGAGTTT 120
QY 2706 TTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCAT 2765
Db |||||
121 TTTAGTGAGGCTGGAAAG-CAATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCAT 179
QY 2766 AAGAGTTCTGTAATCGGAGAACATTATGAATGCACTCTTTTGCACTACCTTCAAGAAAA 2825
Db |||||
180 AAGAGTTCTGTAATCGGAGAACATTATGAATGCACTCTTTTGCACTACCTTCAAGAAAA 239
QY 2826 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 2885
Db |||||
240 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 299
QY 2886 GGTATACAC 2894
Db |||||
300 GGTATACAC 308

RESULT 55
BD109703
LOCUS EST and encoded human protein. 308 bp DNA linear PAT 18-SEP-2002
DEFINITION BD109703
ACCESSION BD109703
VERSION BD109703.1 GI:23204521
KEYWORDS JP 2002010789-A/1780.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo;
1 (bases 1 to 308)
REFERENCE Edwards J.B.D.M., Jobert, S. and Giordano, J.E.
EST and encoded human protein
Patent: JP 2002010789-A 1780 15-JAN-2002;
GENSET CORP

COMMENT
OS Homo sapiens (human)
PN JP 2002010789-A/1780
PD 15-JAN-2002
PF 07-AUG-2000 JP 2002080989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC Von Heijne matrix
CC score 4.5999990463257
CC seq ANTSILSLFLVRA/GK
FH Key Location/Qualifiers
FT CDS 17..178
FT sig_peptide 17..133.
FT Location/Qualifiers
1..308

FEATURES
source
1..308
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 6.0%; Score 188; DB 2; Length 308;
Best Local Similarity 99.7%; Pred. No. 5.4e-98;
Matches 308; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2586 CTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTACTGCT 2645
Db |||||
1 CTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTACTGCT 60
QY 2646 CTTACATGGTTCTCGGATGAGATGTCATTTTGACATPACCAAGTATATTACTGAGTTT 2705
Db |||||
61 CTTACATGGTTCTCGGATGAGATGTCATTTTGACATPACCAAGTATATTACTGAGTTT 120

QY 2706 TTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCAT 2765
Db |||||
121 TTTAGTGAGGCTGGAAAG-CAATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCAT 179
QY 2766 AAGAGTTCTGTAATCGGAGAACATTATGAATGCACTCTTTTGCACTACCTTCAAGAAAA 2825
Db |||||
180 AAGAGTTCTGTAATCGGAGAACATTATGAATGCACTCTTTTGCACTACCTTCAAGAAAA 239
QY 2826 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 2885
Db |||||
240 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 299
QY 2886 GGTATACAC 2894
Db |||||
300 GGTATACAC 308

RESULT 56
AR414150
LOCUS Sequence 1787 from patent US 6639063. 308 bp DNA linear PAT 18-DEC-2003
DEFINITION AR414150
ACCESSION AR414150
VERSION AR414150.1 GI:40169260
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 308)
AUTHORS Edwards, J.-B.D.M., Jobert, S. and Giordano, J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 1787 28-OCT-2003;
Genset S.A.;;
WOX;

FEATURES
source
1..308
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 6.0%; Score 188; DB 2; Length 308;
Best Local Similarity 99.7%; Pred. No. 5.4e-98;
Matches 308; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2586 CTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTACTGCT 2645
Db |||||
1 CTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTACTGCT 60
QY 2646 CTTACATGGTTCTCGGATGAGATGTCATTTTGACATPACCAAGTATATTACTGAGTTT 2705
Db |||||
61 CTTACATGGTTCTCGGATGAGATGTCATTTTGACATPACCAAGTATATTACTGAGTTT 120
QY 2706 TTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCAT 2765
Db |||||
121 TTTAGTGAGGCTGGAAAG-CAATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCAT 179
QY 2766 AAGAGTTCTGTAATCGGAGAACATTATGAATGCACTCTTTTGCACTACCTTCAAGAAAA 2825
Db |||||
180 AAGAGTTCTGTAATCGGAGAACATTATGAATGCACTCTTTTGCACTACCTTCAAGAAAA 239
QY 2826 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 2885
Db |||||
240 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 299
QY 2886 GGTATACAC 2894
Db |||||
300 GGTATACAC 308

RESULT 57
AX522584/c
LOCUS Sequence 254 from Patent WO02064731. 168 bp DNA linear PAT 24-OCT-2002
DEFINITION AX522584
Sequence 254 from Patent WO02064731.

AX522584
 AX522584.1 GI:24411538
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

1
 Telerman, A., Anson, R., Tuijinder, M. and Susini, L.
 Sequences involved in phenomena of tumour suppression, tumour
 reversion, apoptosis and/or virus resistance and their use as
 medicines

JOURNAL
 Patent: WO 02064731-A 254 22-AUG-2002;
 Molecular Engines Laboratories (FR)
 Location/Qualifiers

FEATURES
 source
 1..168
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 5.4%; Score 167; DB 2; Length 168;
 Best Local Similarity 100.0%; Pred. No. 1.3e-85;
 Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2944 GATCATCACATTTTGTGATCTGCGCATGTAACATCTACTCTGAAATAAATGTGTGCA 3003
 |||||
 168 GATCATCACATTTTGTGATCTGCGCATGTAACATCTACTCTGAAATAAATGTGTGCA 109
 |||||

QY 3004 TGCAGGGCTACGGTTTGTGGTAGTAATCTAATCACTTAACCCACATGCTCAAAATCA 3063
 |||||
 108 TGCAGGGCTACGGTTTGTGGTAGTAATCTAATCACTTAACCCACATGCTCAAAATCA 49
 |||||

QY 3064 AATGATACATATTCTCTGAGAGACCCAGCAATACCATAGAAATTACTA 3110
 |||||
 48 AATGATACATATTCTCTGAGAGACCCAGCAATACCATAGAAATTACTA 2
 |||||

RESULT 58
 AC068507/c
 LOCUS
 DEFINITION Homo sapiens chromosome 15 clone RP11-349H15 map 15, WORKING DRAFT
 SEQUENCE, 27 unorderes pieces.
 AC068507
 AC068507.2 GI:8389522
 HTG; HTGS PHASE1; HTGS_DRAFT.
 KEYWORDS
 Homo sapiens (human)
 SOURCE
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 1 (bases 1 to 176282)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,
 Boguslavskiy, L., Bouckhgalter, B., Brown, A., Burkett, G.,
 Campopiano, A., Castelle, A., Choe, P., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,
 Duggan, S., Dono, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 Melidrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,

```

* 55556 55655: gap of 100 bp
* 55656 60496: contig of 4841 bp in length
* 60497 60596: gap of 100 bp
* 60597 68544: contig of 7948 bp in length
* 68545 76940: gap of 100 bp
* 76941 77041: gap of 100 bp
* 77042 85480: contig of 8440 bp in length
* 85481 85580: gap of 100 bp
* 85581 93905: contig of 8325 bp in length
* 93906 94005: gap of 100 bp
* 94006 108063: contig of 14058 bp in length
* 108064 109163: gap of 100 bp
* 109164 123596: contig of 15333 bp in length
* 123597 123597: gap of 100 bp
* 123597 13965: contig of 16369 bp in length
* 13966 140065: gap of 100 bp
* 140066 176282: contig of 36217 bp in length.
FEATURES
    source
    1..176282
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="15"
        /map="15"
        /clone_lib="RPCI-11 Human Male BAC"
        /clone="RP11-349H15"
        /notes="assembly_fragment"
    1..1477
        /estimated_length=100
        /notes="assembly_fragment"
    1578..2785
        /estimated_length=100
        /notes="assembly_fragment"
    2786..2885
        /estimated_length=100
        /notes="assembly_fragment"
    2886..4859
        /estimated_length=100
        /notes="assembly_fragment"
    4860..4959
        /estimated_length=100
        /notes="assembly_fragment"
    4960..6416
        /estimated_length=100
        /notes="assembly_fragment"
    6417..6516
        /estimated_length=100
        /notes="assembly_fragment"
    6517..8739
        /estimated_length=100
        /notes="assembly_fragment"
    8740..8839
        /estimated_length=100
        /notes="assembly_fragment"
    8840..10868
        /estimated_length=100
        /notes="assembly_fragment"
    10869..120098
        /estimated_length=100
        /notes="assembly_fragment"
    120099..13737
        /estimated_length=100
        /notes="assembly_fragment"
    13738..13937
        /estimated_length=100
        /notes="assembly_fragment"
    13938..15329
        /estimated_length=100
        /notes="assembly_fragment"
    15330..15429
        /estimated_length=100
        /notes="assembly_fragment"
    15430..17938
        /estimated_length=100
        /notes="assembly_fragment"
    17939..18038
        /estimated_length=100
        /notes="assembly_fragment"
    18039..19981
        /estimated_length=100
        /notes="assembly_fragment"
    19982..20081
        /estimated_length=100
        /notes="assembly_fragment"
    20082..23587
        /estimated_length=100
        /notes="assembly_fragment"
    23588..23687
        /estimated_length=100
        /notes="assembly_fragment"
    23688..27447
        /estimated_length=100
        /notes="assembly_fragment"

```

```

gap
    27448..27547
        /estimated_length=100
        /note="assembly_fragment"
    27548..30161
        /note="assembly_fragment"
        /vector_side:"left"
    30162..30261
        /estimated_length=100
        /note="assembly_fragment"
    30262..34737
        /estimated_length=100
        /note="assembly_fragment"
    34738..34837
        /estimated_length=100
        /note="assembly_fragment"
    34838..40101
        /estimated_length=100
        /note="assembly_fragment"
    40102..40201
        /estimated_length=100
        /note="assembly_fragment"
    40202..44039
        /estimated_length=100
        /note="assembly_fragment"
    44040..44139
        /estimated_length=100
        /note="assembly_fragment"
    44140..44950
        /estimated_length=100
        /note="assembly_fragment"
    44951..44955
        /estimated_length=100
        /note="assembly_fragment"

Query Match
Best Local Similarity 5.0%; Score 156; DB 12; Length 176282;
Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2331 GGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA 2390
Db 44750 GGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA 44691

QY 2391 TTTCAATTGACTTAGATCGTGTGGCATCCACGCTGGTCTCTATGGAGGATACCTCTCCCT 2450
Db 44690 TTTCAATTGACTTAGATCGTGTGGCATCCACGCTGGTCTCTATGGAGGATACCTCTCCCT 44631

QY 2451 GATGCGATTAAATGACAGAGGTGAGATATCTTCAGGGT 2486
Db 44630 GATGCGATTAAATGACAGAGGTGAGATATCTTCAGGGT 44595

RESULT 59
BD036103
LOCUS
DEFINITION
ACCESSION BD036103
VERSION
KEYWORDS JP 2001269182-A/12349.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 89)
AUTHORS Edwards J.B.D.M., Duclair, E. and Jordan, J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 12349 02-OCT-2001;
GENSET
COMMENT
OS Homo sapiens (human)
PN JP 2001269182-A/12349
PD 02-OCT-2001
PP 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS EIMERIC DUCLAIR, JEAN YVES
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
CC G06F15/40
FH Key Location/Qualifiers.
FEATURES
    source
    1..89
        /organism="Homo sapiens"

```

/db_xref="taxon:9606"		/db_xref="taxon:9606"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity		Best Local Similarity	
Matches		Matches	
2.6%; Score 82; DB 2; Length 89;		2.6%; Score 82; DB 2; Length 89;	
100.0%; Pred. No. 1.9e-35;		100.0%; Pred. No. 1.9e-35;	
0; Mismatches 0; Indels 0; Gaps 0;		0; Mismatches 0; Indels 0; Gaps 0;	
3039 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 3098		3039 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 3098	
1 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 60		1 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 60	
QY		QY	
DB		DB	
3099 TAAGAATTACTAAAAA 3120		3099 TAAGAATTACTAAAAA 3120	
1 TAAGAATTACTAAAAA 82		1 TAAGAATTACTAAAAA 82	
QY		QY	
DB		DB	
RESULT 62		RESULT 62	
BV448370		BV448370	
LOCUS		LOCUS	
DEFINITION		DEFINITION	
Accession		Accession	
Version		Version	
Keywords		Keywords	
Source		Source	
Organism		Organism	
Reference		Reference	
Authors		Authors	
Title		Title	
Journal		Journal	
Features		Features	
Source		Source	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity		Best Local Similarity	
Matches		Matches	
2.6%; Score 82; DB 2; Length 89;		2.6%; Score 82; DB 2; Length 89;	
100.0%; Pred. No. 1.9e-35;		100.0%; Pred. No. 1.9e-35;	
0; Mismatches 0; Indels 0; Gaps 0;		0; Mismatches 0; Indels 0; Gaps 0;	
3039 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 3098		3039 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 3098	
1 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 60		1 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 60	
QY		QY	
DB		DB	
3099 TAAGAATTACTAAAAA 3120		3099 TAAGAATTACTAAAAA 3120	
1 TAAGAATTACTAAAAA 82		1 TAAGAATTACTAAAAA 82	
QY		QY	
DB		DB	
RESULT 61		RESULT 61	
AX900570		AX900570	
LOCUS		LOCUS	
DEFINITION		DEFINITION	
Accession		Accession	
Version		Version	
Keywords		Keywords	
Source		Source	
Organism		Organism	
Reference		Reference	
Authors		Authors	
Title		Title	
Journal		Journal	
Features		Features	
Source		Source	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity		Best Local Similarity	
Matches		Matches	
2.6%; Score 82; DB 2; Length 89;		2.6%; Score 82; DB 2; Length 89;	
100.0%; Pred. No. 1.9e-35;		100.0%; Pred. No. 1.9e-35;	
0; Mismatches 0; Indels 0; Gaps 0;		0; Mismatches 0; Indels 0; Gaps 0;	
3039 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 3098		3039 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 3098	
1 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 60		1 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 60	
QY		QY	
DB		DB	
3099 TAAGAATTACTAAAAA 3120		3099 TAAGAATTACTAAAAA 3120	
1 TAAGAATTACTAAAAA 82		1 TAAGAATTACTAAAAA 82	
QY		QY	
DB		DB	
Protocol:		Protocol:	
Hot Start:		Hot Start:	
Denaturation:		Denaturation:	
Annealing:		Annealing:	
Polymerization:		Polymerization:	
PCR Cycles:		PCR Cycles:	
Extension:		Extension:	
Thermal Cycler:		Thermal Cycler:	
Template:		Template:	
Primer:		Primer:	
dNTP's:		dNTP's:	
Tag Polymerase:		Tag Polymerase:	
Fidelity, Roche:		Fidelity, Roche:	
Total Vol:		Total Vol:	
Buffer:		Buffer:	
MgCl2:		MgCl2:	
Fast Start polymerase reaction buffer (Roche)		Fast Start polymerase reaction buffer (Roche)	
Bases 1-320 are 97% homologous (Blast) to bases 2805-3124 of		Bases 1-320 are 97% homologous (Blast) to bases 2805-3124 of	
NM_197961.1. Primers were chosen to amplify genomic DNA in the 3'		NM_197961.1. Primers were chosen to amplify genomic DNA in the 3'	
region of DPP8. As human sequence was used to design the primers,		region of DPP8. As human sequence was used to design the primers,	
the primer sequences are not included in the rhesus sequence		the primer sequences are not included in the rhesus sequence	
provided below. To obtain additional information regarding primers		provided below. To obtain additional information regarding primers	
or clones contact: Dr. Robert Norgren; Dept of Genetics, Cell		or clones contact: Dr. Robert Norgren; Dept of Genetics, Cell	

A database containing sequences associated with this project can be found at: <http://rhesusgenechip.unomaha.edu/index.html>

```

/organism="Macaca mulatta"
/mol_type="genomic DNA"
/strain="Indian origin"
/db_xref="taxon:9554"
/clone="WM9839"
/clone_lib="Rhesus macaque genomic DNA"
/dev_stage="Adult"
/note="Organ: Liver; Vector: pGEM-T Easy
STS was amplified from rhesus genomic DNA
forward and reverse primers listed above
into pGEM-T Easy"
1. .477
/gene="pppg"
/note="dipeptidylpeptidase 8"
<1. .477
/gene="pppg"

```

Matches	82;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

214 GTTTGGTGTAGTAATCTAATACTTAACCCACATGCTCAAAATCAAATGATACATATTC 273

274 CTGAGAGACCCAGCAATACCAT 295

AC157219
CUS
157219

AC157219
AC157219 2 CT-68001506

ORGANISM	Metazoa: Chordata
<i>Bos taurus</i> (cattle)	

REFERENCE
AUTHORS
1 (bases 1 to 222034)
Muzny D Maria Metabolite
1 (bases 1 to 222034)
Muzny D Maria Metabolite

Baldwin, D., Bandaranaika, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K. and Blair, T. 1997. *Journal of the American Veterinary Medical Association*, 261, 1031-1034.

Sardinas, V., Carter, R., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Cleveland, C., Cochrane, S., Cochrane, T., Chu, J.,

Sergueev, O., Benson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Facotto, M., Farnsworth, C.

Garcia, A., Garner, T., Garza, M.,
Garner, T., Garza, M., Garza, W.,
Gebegeorgis, E., Geer, K., Gill, R.,
Grady, M., Guerra, W., Guevara, W.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idiebird, D. Jackson, A.,
Jackson I., Jacobson, R., Jones, K., Laidlaw, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idiebird, D. Jackson, A.,

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levay, J., Lewals, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Loresnushewa, L., Louisseghed, H., Lozard, R.J., Lu, X., Ma, J., Maneshwar, M., Mahindran, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapus, P., Martin, K., Martin, R., Martin, E., Mathew, S., McLeod, M.P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankweli, O., Okwunou, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, W.A., Reigh, R., Reilly, B., Reilly, J., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scheer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smaja, D., Sneed, A., Sodergren, E., Song, X.-Z., Soresler, R., Soosa, J., Steimle, M., Strong, R., Sutton, A., Svacek, A., Taber, P., Taylor, C., Taylor, T., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasaña, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weis, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

REFERENCE

1

TITLE
JOURNAL

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 29449: contig of 29449 bp in length
* 29450 29615: gap of 166 bp
* 29616 35158: contig of 5543 bp in length
* 35159 35208: gap of 50 bp
* 35209 46041: contig of 10833 bp in length
* 46042 46091: gap of 50 bp
* 46092 65722: contig of 19630 bp in length
* 65723 65771: gap of 50 bp
* 65772 70033: contig of 4262 bp in length
* 70034 70084: gap of 50 bp
* 70085 75271: contig of 5187 bp in length
* 75272 75459: gap of 189 bp
* 75460 77511: contig of 2051 bp in length
* 77512 77739: gap of 229 bp
* 77740 87986: contig of 10247 bp in length
* 87987 88036: gap of 50 bp
* 88037 89277: contig of 1891 bp in length
* 89278 89977: gap of 50 bp
* 89978 92059: contig of 2082 bp in length
* 92060 92109: gap of 50 bp
* 92110 93888: contig of 1779 bp in length
* 93889 93939: gap of 50 bp
* 93940 98457: contig of 4518 bp in length
* 98458 98506: gap of 50 bp
* 98507 115454: contig of 16948 bp in length
* 115455 115504: gap of 50 bp
* 115505 116861: contig of 1357 bp in length
* 116862 116911: gap of 50 bp
* 116912 129154: contig of 12243 bp in length
* 129155 129204: gap of 50 bp
* 129205 137025: contig of 7820 bp in length
* 137026 137074: gap of 50 bp
* 137075 156249: contig of 19175 bp in length
* 156250 156299: gap of 50 bp
* 156300 159047: contig of 2747 bp in length
* 159048 159096: gap of 50 bp
* 159097 165860: contig of 6764 bp in length
* 165861 165910: gap of 50 bp
* 165911 178253: contig of 12343 bp in length
* 178254 178303: gap of 50 bp
* 178304 180139: contig of 1896 bp in length
* 180140 180249: gap of 50 bp
* 180250 217650: contig of 37401 bp in length
* 217651 217750: gap of unknown length
* 217751 218931: contig of 1181 bp in length
* 218932 219031: gap of unknown length
* 219032 220396: contig of 1365 bp in length
* 220397 220436: gap of unknown length
* 220437 222034: contig of 1538 bp in length.

FEATURES

source

1. 222034
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-74L16"
29450..29615
/estimated_length=166
35159..35208
/estimated_length=50
46042..46091
/estimated_length=50
65722..65771
/estimated_length=50
70034..70083
/estimated_length=50

gap
gap
gap
gap
gap

75271..75459
/estimated_length=189
77511..77739
/estimated_length=229
87987..88036
/estimated_length=50
89928..89977
/estimated_length=50
92050..92109
/estimated_length=50
93889..93938
/estimated_length=50
98457..98506
/estimated_length=50
115455..115504
/estimated_length=50
116862..116911
/estimated_length=50
129155..129204
/estimated_length=50
137025..137074

Query Match 2.6%; Score 81; DB 12; Length 222034;
Best Local Similarity 100.0%; Pred.No. 8.8e-35;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1340 CCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTTCATCTACCTGAATTAT 1399
Db 118848 CCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTTCATCTACCTGAATTAT 118907
QY 1400 TTATCCAGTAGAAGATGATG 1420
Db 118908 TTATCCAGTAGAAGATGATG 118928

RESULT 64
AC106509/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-20218, WORKING DRAFT SEQUENCE, 3
unordered pieces.
AC106509
AC106509.4 GI:30580817
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 224227)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Geer,K., Gill,R., Grady,M., Garcia,A., Garner,T., Garza,M.,
Gebregioris,E., Geor,K., Hamill,C., Hamilton,C., Hernandez,J.,
Harvey,Y., Havlak,P., Hawes,A., Hladun,S.L., Hodgson,A., Hughes,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

224227 bp DNA linear HTG 13-MAY-2003
Rattus norvegicus clone CH230-20218, WORKING DRAFT SEQUENCE, 3
unordered pieces.

AC106509
AC106509.4 GI:30580817
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)

ORGANISM
Rattus norvegicus
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 224227)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Geer,K., Gill,R., Grady,M., Garcia,A., Garner,T., Garza,M.,
Gebregioris,E., Geor,K., Hamill,C., Hamilton,C., Hernandez,J.,
Harvey,Y., Havlak,P., Hawes,A., Hladun,S.L., Hodgson,A., Hughes,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

AC106509
AC106509.4 GI:30580817
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)

ORGANISM
Rattus norvegicus
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 224227)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Geer,K., Gill,R., Grady,M., Garcia,A., Garner,T., Garza,M.,
Gebregioris,E., Geor,K., Hamill,C., Hamilton,C., Hernandez,J.,
Harvey,Y., Havlak,P., Hawes,A., Hladun,S.L., Hodgson,A., Hughes,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

AC106509
AC106509.4 GI:30580817
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)

ORGANISM
Rattus norvegicus
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 224227)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Geer,K., Gill,R., Grady,M., Garcia,A., Garner,T., Garza,M.,
Gebregioris,E., Geor,K., Hamill,C., Hamilton,C., Hernandez,J.,
Harvey,Y., Havlak,P., Hawes,A., Hladun,S.L., Hodgson,A., Hughes,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

AC106509
AC106509.4 GI:30580817
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)

ORGANISM
Rattus norvegicus
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

Lorensuheva, L., Loulsegged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapus, P., Martin, K., Martin, R., Martinez, E., Mathiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelumb, O., Okuwonu, G., Olarunpasegoun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Ploppan, F., Poindestre, A., Popovic, D., Primus, E., Pu, L.-L., Puccio, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rokey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Sreemle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

Direct Submission
Unpublished
2 (bases 1 to 224227)
Worley, K. C.

Submitted (12-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 224227)

REFERENCE AUTHORS TITLE JOURNAL

Rat Genome Sequencing Consortium.

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:22857495. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLBA
Center clone name: CH230-20218
----- Summary Statistics

Assembly program: Atlas 3.0;
Consensus quality: 211258 bases at least Q30
Consensus quality: 214044 bases at least Q40
Consensus quality: 215992 bases at least Q20
Estimated insert size: 220399; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 221791: contig of 221791 bp in length
* 221792 221891: gap of unknown length
* 221892 222903: contig of 1018 bp in length
* 222910 224227: gap of unknown length
* 223010 224227: contig of 1218 bp in length.

FEATURES

source

1..224227
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clones="CH230-20218"

misc_feature

1..1752
/note="wgs_end_extension
clone_end:Sp6"
3617..3835
/note="clone_boundary
clone_end:Sp6
site:ECORI
end_sequence:BH334357"
220120..221791
/note="wgs_contig"
221792..221891
/estimated_length=unknown
222910..223009
/estimated_length=unknown

misc_feature

220120..221791
/note="wgs_contig"
221792..221891
/estimated_length=unknown
222910..223009
/estimated_length=unknown

gap

220120..221791
/note="wgs_contig"
221792..221891
/estimated_length=unknown
222910..223009
/estimated_length=unknown

gap

220120..221791
/note="wgs_contig"
221792..221891
/estimated_length=unknown
222910..223009
/estimated_length=unknown

ORIGIN

Query Match

Best Local Similarity 2.0%; Score 62; DB 12; Length 224227;

Matches

62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,U.,
 Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 4799)
 Direct Submission
 Submitted (09-JAN-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
 Thomas L. Casavant.
 Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
 Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
 Scheetz,T., Smith,C., Shier,E., Tack,D., Trout,K., Walters,J.,
 Casavant,T., Soares,M.B.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: Plate: Row: Column: 0
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 27229207.

FEATURES

source

1. .4799
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="MGC:58057 IMAGE:6410075"
 /tissue_type="Brain, enriched mouse brain 12.5dp"
 /clone_lib="NIH BMAP_F00"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"
 1. .4799
 /gene="Dpp8"
 /db_xref="GeneID:74388"
 /db_xref="MGI:1921638"
 288. .2966
 /gene="Dpp8"
 /codon_start=1
 /product="dipeptidylpeptidase 8"
 /protein_id="AA43124.1"
 /db_xref="GI:27695450"
 /db_xref="GeneID:74388"
 /db_xref="MGI:1921638"

gene

CDS

ORIGIN
 Query Match 1.8%; Score 56; DB 6; Length 4799;
 Best Local Similarity 100.0%; Pred. No. 4.6e-20;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1999 AAAACAAGGAATTTTGGCCACCAATTTTGATTCAGCAGTCTCTTCTGACTA 2054
 |||||
 Db 2103 AAAACAAGGAATTTTGGCCACCAATTTTGATTCAGCAGTCTCTTCTGACTA 2158
 |||||
 RESULT 66
 LOCUS BC059222
 DEFINITION Mus musculus dipeptidylpeptidase 8, mRNA (cDNA clone MGC:66620
 IMAGE:6410075), complete cds.
 ACCESSION BC059222.1 GI:37590653
 VERSION BC059222
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.
 1 (bases 1 to 4799)
 Strausberg,R.B., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heideh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shvachenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,U.,
 Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 4799)
 Strausberg,R.
 Direct Submission
 Submitted (01-OCT-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 125 Row: i Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 27229207.

FEATURES
source

Location/Qualifiers
1..4799
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:66620 IMAGE:6410075"
/tissue_type="brain, enriched mouse brain 12.5dp"
/clone_lib="NIH BMAP_F00"
/lab_host="DH10B"
/note="vector: pYX-ASC"

gene

1..4799
/gene="Dpp8"
/db_xref="GeneID:74388"
/db_xref="WGI:1921638"
288..2966
/gene="Dpp8"
/codon_start=1
/product="dipeptidylpeptidase 8"
/protein_id="AAH59222.1"
/db_xref="GI:137590654"
/db_xref="GeneID:74388"
/db_xref="WGI:1921638"

CDS

translation="MKIPGRCNMAAMETQLGVETPETAECGEGSQDRPKLEP
FYERSWQLKKLADTRKHGVMMAKAPHDFMVKRTDPDGPHSDRVVYLA
KENTLYSEPTPTINRAAVILMSWPLDLFQATLDYGMVSREELRLRKRIGT
AAYDHPGSGTFLPQAGSGIYHIKDGPHGFTQQLPLNLVETSCPNRMWPKLCAD
PMWIAFHNDLWISNLVTRERTIYVHNLANNEEDPRSGAVATPVLQERDYS
YMWCPQARTPSGGKILRLYLEENDESEVEIHTVSPMLERADSPVFKTGNPK
VPMSEIVVDAAGGIIDVIDKELVQPEILFVGSVYIARAGWTPEGKAWSLLDRS
QTHQLVLSPLEFLPVEDDMDRLRIEVPDSVTPILIIYEETDIDWNIHDPIHV
TQHEDEIFPASCCKTGFRHLVKITSLKESYKRSSGGLPAPDKCPKLTKEITI
HSCCLSRHCDFFISYKNQPHCVSLYKLSPPEDDPVHKTSFNATILDSAGPLDY
TPPEIFSTFTGFLYGLMKPHDLOPKKYPTVLYFGPOVLVNNRFGKVKFYR
LNTLASLGVVVVINDRSGHRLGFKFAGKYKGQLEIDVDVEGLYLASQDFIDL
DRVGHGYSYGVISLMALMORSDFRVALAGPVTMLIEIDVGYTERNGHPDQEQ
GYLGSVAQAEKFPSEPNLLHLGLDENVHAHTSILLSFLVAGAPYDILQIYPO
EHRISVPESGSEHYELHLLVYLQENLSRAALKYI"

ORIGIN

Query Match 1.8%; Score 56; DB 6; Length 4799;
Best Local Similarity 100.0%; Pred. No. 4.6e-20;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1999 AAAACAAGGAATTTTGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTA 2054
|||||
Db 2103 AAAACAAGGAATTTTGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTA 2158
|||||

RESULT 67

AC144143/c
LOCUS AC144143 305843 bp DNA linear HTG 09-APR-2003
DEFINITION Macaca mulatta clone CH250-272H19, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC144143
VERSION AC144143.1 GI:29649813
KEYWORDS HTG; HTGS PHASE2; HTGS_FGI.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta

REFERENCE

1. (bases 1 to 305843)
2. (bases 1 to 305843)
AUTHORS Ceuros,M. and Milosavljevic,A.
TITLES Pooled genomic indexing (PGI): mathematical analysis and experiment design
JOURNAL (in) Guigo,R. and Gusfield,D. (Eds.);
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
Springer (2002)
REFERENCE 2 (bases 1 to 305843)
AUTHORS Milosavljevic,A., Sodergren,E., Ceuros,M., Li,B., Jackson,A.R.,
Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
3 (bases 1 to 305843)
Worley,K.C.
Direct Submission
Submitted (09-APR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project information
Center project name: LCRF
Center clone name: CH250-272H19
----- Summary Statistics
Chemistry: Dye-terminator Big Dye; Inf% of reads
Chemistry: Dye-terminator Big Dye; Inf% of reads
Consensus quality: 12640 bases at least Q40
Consensus quality: 15085 bases at least Q30
Consensus quality: 17385 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: The contigs are based on the application
* of the PGI method using the Human genome (NCBI build 31)
* as the comparative genome.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have

COMMENT


```

* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 305843: contig of 305843 bp in length.
FEATURES
    source
        1..305843
            /organism="Macaca mulatta"
            /mol_type="genomic DNA"
            /db_xref="taxon:9544"
            /clone="CH250-272H19"
            /note="assembly name:CH250-272H19.1B
            CONFIDENCE: 0.83"
    misc_feature
        1..305843
            /note="assembly name:CH250-272H19.1B
            CONFIDENCE: 0.83"
ORIGIN
Query Match          1.7%; Score 54; DB 12; Length 305843;
Best Local Similarity 100.0%; Pred. No. 7.6e-19;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1167 AGGTACAGCAATCTTAAGTCACCTTTTAAGATGTCAGAAATATGATTGATGC 1220
      |||
Db 131782 AGGTACAGCAATCTTAAGTCACCTTTTAAGATGTCAGAAATATGATTGATGC 131729

RESULT 68
AC144143 305843 bp DNA linear HTG 09-APR-2003
LOCUS Macaca mulatta clone CH250-272H19, *** SEQUENCING IN PROGRESS ***
AC144143
AC144143.1 GI:29649813
HTG: HTGS PHASE2; HTGS PGI.
Macaca mulatta (rhesus monkey)
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.
ORGANISM
1 (bases 1 to 305843)
Cauros,M. and Milosavljevic,A.
Pooled genomic indexing (PGI): mathematical analysis and experiment
design
(in) Guigo,R. and Gusfield,D. (Eds.);
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
Springer (2002)
2 (bases 1 to 305843)
Milosavljevic,A., Sodergren,E., Cauros,M., Li,B., Jackson,A.R.,
Adams,C., Adio-Oduola,B., Ali-ouman,F.R., Allen,C., Alsbrooks,S.L.,
Amaratunga,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Dox,C., Coyle,M.D., Dathorne,S.R., David,K.R., Deigado,O.,
Davy-Garroll,L., Dederich,D.A., Delaney,K.J., Draper,H.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Edwards,C.C.,
Eugan-Rocha,S., Durbin,K.J., Egan,A., Earnhart,C., Edwards,C.C.,
Dulanja-C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
Elhaj,C., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Flagg,N., Ford,J., Foster,P., Frantz,P., Gill,R., Gorrell,J.H., Guevara,W.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Harris,K.,
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Hernandez,O.,
Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hollins,B., Homsif.,
Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Ioshikhes,I., Jackson,L.E.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Joudah,S.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Korvah,J., Kovar,C.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Lee,E., Lewis,L.C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Liu,W.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Lucier,R., Lina,R.,
Louiseged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Martin,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondel,I., McLeod,M.P.,
Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,

```

```

Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,M., Neal,D., Neilson,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
Ogih,M., Okwona,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
Shooshtari,N., Sisson,I., Sodergren,E., Sonaik,H., Spark,A.,
Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Williams,R.,
Washington,C., Watlington,S., Williams,G., Williamson,A.,
Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
3 (bases 1 to 305843)
Worley,K.C.
Direct Submission
Submitted (09-APR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: LCRF
Center clone name: CH250-272H19
----- Summary Statistics
Chemistry: Dye-terminator Big Dye: inf% of reads
Chemistry: Dye-terminator Big Dye: inf% of reads
Consensus quality: 12640 bases at least Q40
Consensus quality: 15085 bases at least Q30
Consensus quality: 17385 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: The contigs are based on the application
* of the PGI method using the Human genome (NCBI build 31)
* as the comparative genome.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 305843: contig of 305843 bp in length.
* Location/Qualifiers
    1..305843
        /organism="Macaca mulatta"
        /mol_type="genomic DNA"
        /db_xref="taxon:9544"
        /clone="CH250-272H19"
    1..305843
        /note="assembly name:CH250-272H19.1B
        CONFIDENCE: 0.83"
    misc_feature
        1..305843
            /note="assembly name:CH250-272H19.1B
            CONFIDENCE: 0.83"
ORIGIN
Query Match          1.5%; Score 48; DB 12; Length 305843;
Best Local Similarity 100.0%; Pred. No. 2.6e-15;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 978 TGTTCTCCAGAGAAATTTGATAGATATTCCTGGCTATTGGTGTGTC 1025
      |||
Db 249937 TGTTCTCCAGAGAAATTTGATAGATATTCCTGGCTATTGGTGTGTC 249984

```

RESULT 69

AC161443
 LOCUS Mus musculus chromosome 1, clone RP23-404K2, complete sequence.
 DEFINITION AC161443
 VERSION AC161443.14 GI:82177432
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 186144)

AUTHORS

Birken, B., Nusbaum, C. and Lander, E.

TITLE

Mus musculus chromosome 1, clone RP23-404K2

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 186144)

AUTHORS

Birken, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,

Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
 Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
 DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
 Galagan, J., Gardyna, S., Graham, L., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
 MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemбек, L.,
 Zimmer, A. and Zody, M.

Direct Submission

Submitted (14-MAY-2005) Broad Institute of MIT and Harvard, 320

JOURNAL

Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 186144)

AUTHORS

Birken, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,

Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
 Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
 DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
 Galagan, J., Gardyna, S., Graham, L., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
 MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemбек, L.,
 Zimmer, A. and Zody, M.

Direct Submission

Submitted (21-OCT-2005) Broad Institute of MIT and Harvard, 320

JOURNAL

Charles Street, Cambridge, MA 02141, USA

REFERENCE

4 (bases 1 to 186144)

AUTHORS

Birken, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,

Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
 Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
 DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,

Galagan, J., Gardyna, S., Graham, L., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
 MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemбек, L.,
 Zimmer, A. and Zody, M.

Direct Submission
 Submitted (12-NOV-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

On Nov 12, 2005 this sequence version replaced gi:77798131.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L32392

Center clone name: 404_K_2

----- Location/Qualifiers

1. 186144

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="1"

/map="1"

/clone="RP23-404K2"

/clone_lib="RP23-404K2"

1. 272

/rpt_family="L1Md_F2"

273. 1268

/rpt_family="L1Md_F2"

complement(1428..1880)

/rpt_family="RLR17"

complement(2701..2775)

/rpt_family="L1Md_F2"

3013..3045

/rpt_family="HAL1b"

3046..3070

/rpt_family="TTTGTG n"

3071..3133

/rpt_family="HAL1b"

complement(4980..5501)

/rpt_family="Lx2"

5691..5713

/rpt_family="TTTGTG n"

6467..6493

/rpt_family="(A) n"

6541..6866

/rpt_family="ORR1A3"

7774..7807

/rpt_family="AT_Rich"

8071..8211

/rpt_family="BLF"

8656..8732

/rpt_family="(GA) n"

complement(8970..9532)

/rpt_family="RLR31_Mm"

9548..9602

/rpt_family="(TTTTTG) n"

complement(9603..9750)

```
/rpt_family="B1_Musl1"  
10432..10616  
/rpt_family="(GGGAGA)n"  
complement(10666..10880)  
/rpt_family="B3"  
complement(11044..11107)  
/rpt_family="L2"  
11488..11664  
/rpt_family="MT2C_Mm"  
complement(11665..13640)  
/rpt_family="MusHAL1"  
complement(13641..13741)  
/rpt_family="L1_Musl1"  
complement(13742..14913)  
/rpt_family="MusHAL1"  
14921..15157  
/rpt_family="RLTR31_Mm"  
complement(15338..15392)  
/rpt_family="MusHAL1"  
complement(15392..16175)  
/rpt_family="L1VL2"  
16186..16471  
/rpt_family="RLTRL1A"  
16472..16508  
/rpt_family="(TG)n"  
16509..16581  
/rpt_family="RLTRL1A"  
16587..16655  
/rpt_family="GA-rich"  
complement(16672..16986)  
/rpt_family="Lx4A"  
complement(16987..17961)  
/rpt_family="RLTRL3A3"  
complement(17962..19206)  
/rpt_family="Lx4A"  
19438..19498  
/rpt_family="(TC)n"  
19499..19539  
/rpt_family="(CA)n"  
19735..19759  
/rpt_family="(TTTA)n"  
complement(19769..20466)  
/rpt_family="Lx4B"  
22311..22333  
/rpt_family="AT_rich"  
22457..22525  
/rpt_family="RLTR40"  
complement(22526..23036)  
/rpt_family="RLTRLB"  
23037..23832  
/rpt_family="RLTR40"  
25559..25637  
/rpt_family="ID4_"  
26627..26791  
/rpt_family="--ORR1A3"  
26919..26958
```

[illegible]

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominiidae; Homo.

REFERENCE   1
AUTHORS     Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.
TITLE       Oligonucleotide library for detecting rna transcripts and splice
            variants that populate a transcriptome
JOURNAL     Patent: WO 0210449-A 15085 07-FEB-2002;
            Compugen Inc. (US)

FEATURES             Location/Qualifiers
                 1...60
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

ORIGIN

Query Match      1.5%; Score 46; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.3e-14;
Matches          46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      832  GCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATTGGA 877
          |||||
          |||||

DB      15  GCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATTGGA 60

RESULT 71
AC171824
LOCUS      AC171824
DEFINITION Bos taurus clone CH240-528L18, *** SEQUENCING IN PROGRESS ***, 15
            linear
            HTG 25-JAN-2006

ACCESSION  AC171824
VERSION    AC171824.2 GI:85702445
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
            Bos taurus (cattle)

```

ORGANISM	REFERENCE AUTHORS
Bos taurus	
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:	
Mammalia: Eutheria: Laurasiatheria: Cetartiodactyla: Ruminantia:	
Pecora: Bovidae: Bovinae: Bos.	
1. (bases 1 to 197073)	
	Muzny, D., Adams, C., Agbai II, O., Allen, C., Alsbrooks, S., Archer, P., Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R., Beraducci, A., Biswal, K., Blyth, P., Bonham, K., Bunay, C., Burch, P., Cadorese, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I., Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, G., Chen, R., Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A., Curry, S., Dai, W., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinu, H., Donlin, J., McCauley, S., Dugan-Kocha, S., Dunn, A., Durbin, K., Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M., Fernandez, S., Fernando, P., Flaggi, N., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T., Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D., Gonzalez-Caray, M., Guevara, W., Holder, M., Haaland, W., Haeblerien, K., Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P., Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J., Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B., Howell, L.T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalatous, K., Kelly, S., Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F., Leal, S., Lee, K., Lee, S., Legall, F.I., Lemon, S., Lewis, L., Li, B., Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y., Liu, Y., Liyanage, D., London, P., Lopez, J., Lorensuheva, L., Lozado, R., Luk, T., Madu, R., Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E., McClelland, H., McPherson, J., Mercadado, C., Metzger, M., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M., Murray, D., Nazarith, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E., Nott, A., Nwaokemele, O., Obregon, M., Ochi-Okorie, C., Odeh, E., Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J., Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,

Rivas, C., Rodriguez, P., Rojas, A., Ruiz, S.J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, Y., Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R., Trejos, Z., Umami, K., Vargo, C., Verduzco, D., Villaana, D., Virk, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 197073)
Worley, K.C.
Direct Submission
Submitted (11-NOV-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 197073)

Bovine Genome Sequencing Consortium
Direct Submission
Submitted (25-JAN-2006) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jan 25, 2006 this sequence version replaced gi:82174753.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: PRK1
Center clone name: CH240-528L18
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 190844 bases at least Q40
Consensus quality: 192231 bases at least Q30
Consensus quality: 193415 bases at least Q20
Estimated insert size: 192666; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 21967: contig of 21967 bp in length
21968 22017: gap of 50 bp
22018 23050: contig of 1033 bp in length
23051 23159: gap of 109 bp
23160 24660: contig of 1501 bp in length
24661 24760: gap of unknown length
24761 26134: contig of 1374 bp in length
26135 26940: gap of 806 bp
26941 38125: contig of 11185 bp in length
38126 38175: gap of 50 bp

Rivers, C., Rodriguez, P., Rojas, A., Ruiz, S.J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, Y., Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R., Trejos, Z., Umami, K., Vargo, C., Verduzco, D., Villaana, D., Virk, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 197073)
Worley, K.C.
Direct Submission
Submitted (11-NOV-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 197073)

Bovine Genome Sequencing Consortium
Direct Submission
Submitted (25-JAN-2006) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jan 25, 2006 this sequence version replaced gi:82174753.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: PRK1
Center clone name: CH240-528L18
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 190844 bases at least Q40
Consensus quality: 192231 bases at least Q30
Consensus quality: 193415 bases at least Q20
Estimated insert size: 192666; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 21967: contig of 21967 bp in length
21968 22017: gap of 50 bp
22018 23050: contig of 1033 bp in length
23051 23159: gap of 109 bp
23160 24660: contig of 1501 bp in length
24661 24760: gap of unknown length
24761 26134: contig of 1374 bp in length
26135 26940: gap of 806 bp
26941 38125: contig of 11185 bp in length
38126 38175: gap of 50 bp

38176 48867: contig of 10692 bp in length
48868 48917: gap of 50 bp
48918 58218: contig of 9300 bp in length
58219 58268: gap of 50 bp
58269 160395: contig of 102128 bp in length
160396 161001: gap of 606 bp
161002 168377: contig of 7376 bp in length
168378 168427: gap of 50 bp
168428 179285: contig of 10857 bp in length
179286 179427: gap of 143 bp
179428 191688: contig of 12260 bp in length
191689 192959: gap of unknown length
192960 193058: contig of 1171 bp in length
193059 194632: gap of unknown length
194633 194731: contig of 1573 bp in length
194732 195949: gap of unknown length
195950 196048: contig of 1217 bp in length
196049 197073: contig of 1025 bp in length.

FEATURES
Location/Qualifiers
1..197073
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clones="CH240-528L18"
21968..22017
/estimated_length=50
23051..23159
/estimated_length=109
24661..24760
/estimated_length=unknown
26135..26940
/estimated_length=806
38126..38175
/estimated_length=50
48868..48917
/estimated_length=50
58218..58267
/estimated_length=50
160396..161001
/estimated_length=606
168378..168427
/estimated_length=50
179285..179427
/estimated_length=143
191688..191787
/estimated_length=unknown
192959..193058
/estimated_length=unknown
194632..194731
/estimated_length=unknown
195949..196048
/estimated_length=unknown

ORIGIN
Query Match 1.4%; Score 44; DB 12; Length 197073;
Best Local Similarity 100.0%; Pred. No. 6e-13;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 271 TGTGAGGAGATATTGATTCACAGATCGCCTAAATTGGAGCC 314
|||||
Db 22514 TGTGAGGAGATATTGATTCACAGATCGCCTAAATTGGAGCC 22557
|||||

RESULT 72
AC112162/c
LOCUS
DEFINITION Mus Musculus chromosome 9 BAC clone MGS1-296M6 ES cell line,
complete sequence.
AC112162
AC112162.8 GI:21327375
VERSION
KEYWORDS
SOURCE Mus musculus (house mouse)

SEQUENCING READ COVERAGE: Attempts are made to complete double stranded sequence for all regions. All sequence is completed to a standard of coverage with a minimum of 3 reads with no ambiguities. If the sequence coverage for a region does not meet this standard, it is indicated in the annotation as low coverage. Low coverage linkages are verified by PCR product size verification or verification of forward and reverse reads from clones which span the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated average error rate is less than 1 per 10,000 bases using the Consed quality parameters. Regions that do not meet this requirement are annotated as Low Quality.

-----Summary Statistics-----
 Center project name: ALI
 Sequencing vector: pSMART; AP399742
 Chemistry: Dye-terminator Big Dye; 100%
 Assembly program: Phrap version 0.990319
 Contig length: 116580
 Fraction of Phrap value < 40: 0
 Error Rate in Consed: 0.00 per 10,000 bases
 Number of N's in Consensus: 0

----- Distribution of Quality < 40 Bases: -----

	1000	900	800	700	600	500	400	300	200	100	0
#											
bases											
	5	10	15	20	25	30	35	40			

	Phrap Value Range										

FEATURES	source	Location/Qualifiers
STS	repeat_region	1. 116580 /organism="Mus musculus" /mol_type="genomic DNA" /strain="ES cell line" /db_xref="taxon:10090" /chromosome="9" /clone="MGS1-296W6" 235. 369 /standard_name="D9Mit316" 258. 303 /rpt_family=" (TG)n" 612. 1204 /standard_name="Nope" complement(1594..1643) /rpt_family="B4A" 1710..1748 /rpt_family=" (TG)n" complement(3793..3909) /rpt_family="PB1D10" 4042..4191 /rpt_family="MIR" 4247..4435 /rpt_family="B2_Mm2" complement(4504..4623) /rpt_family="RSINE1" 4720..4831 /rpt_family="L2" complement(4840..4879) /rpt_family="MIR" 4938..5118 /rpt_family="B2_Mm2"

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.
 1 (bases 1 to 116580)
 Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenccheck,W., Xi,C., Juels,P. and Kucherlapati,R.
 High throughput Mouse Sequencing
 Unpublished
 2 (bases 1 to 116580)
 Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenccheck,W., Xi,C., Juels,P. and Kucherlapati,R.
 Direct Submission
 Submitted (20-FEB-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
 3 (bases 1 to 116580)
 Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenccheck,W., Xi,C., Juels,P. and Kucherlapati,R.
 Direct Submission
 Submitted (04-MAY-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
 4 (bases 1 to 116580)
 Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenccheck,W., Xi,C., Juels,P. and Kucherlapati,R.
 Direct Submission
 Submitted (04-JUN-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
 5 (bases 1 to 116580)
 Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenccheck,W., Xi,C., Juels,P. and Kucherlapati,R.
 Direct Submission
 Submitted (06-JUN-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
 6 (bases 1 to 116580)
 Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenccheck,W., Xi,C., Juels,P. and Kucherlapati,R.
 Direct Submission
 Submitted (21-SEP-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
 On Jun 6, 2002 this sequence version replaced gi:21321767.
 -----Genome Center:
 Center: Harvard Partners Genome Center
 Center Code: HPGC
 Web site: <http://www.hpcgg.org/Sequence/mouse.html>
 Contact: hpgc@mcg.harvard.edu

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
 STSS are identified using ePCR (Genome Res. 7:541-550).
 Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
 Genes and regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

	clone_lib="RPCL-24"
unsure	10180..10254
	/note="Sequence derived from one plasmid subclone."
unsure	51980..52027
	/note="Sequence derived from one plasmid subclone."
unsure	153091..153320
	/note="Unresolved simple sequence repeat."
ORIGIN	
Query Match	1.3%; Score 42; DB 6; Length 155733;
Best Local Similarity	100.0%; Pred.No. 9.le-12;
Matches 42; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1999 AAACAAGGAATTTGGGCCACCATTTTGGATTGAGCAGGT 2040
DB	96576 AAACAAGGAATTTGGGCCACCATTTTGGATTGAGCAGGT 96535
RESULT 74	
AC122463	162528 bp DNA linear ROD 28-JAN-2005
LOCUS	Mus musculus BAC clone RP24-296C19 from 9, complete sequence.
DEFINITION	
AC122463	
VERSION	AC122463.4 GI:50839061
KEYWORDS	HTG.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 162528) Tomlinson,C., Cotton,M. and Haglund,K. The sequence of Mus musculus BAC clone RP24-296C19 Unpublished (2001)
AUTHORS	McPherson,J.D. and Waterston,R.H.
TITLE	Direct Submission
JOURNAL	Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	3 (bases 1 to 162528) Wilson,R.K. Direct Submission Submitted (10-JUL-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
AUTHORS	Wilson,R.K.
TITLE	Direct Submission
JOURNAL	Submitted (30-JUL-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	5 (bases 1 to 162528) Wilson,R.K. Direct Submission Submitted (28-JAN-2005) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
AUTHORS	
TITLE	
JOURNAL	
COMMENT	On Jul 30, 2004 this sequence version replaced gi:50201506. ----- Genome Center Center: Washington University Genome Sequencing Center Center code: WUGSC Web site: http://genome.wustl.edu Contact: submissions@watson.wustl.edu ----- Summary Statistics ----- Center project name: M_BB0296C19 -----
NOTE:	This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence as compressions and repeats; and the assembly was confirmed by

```
repeat_region /rpt_family="ID"
13532..13720
repeat_region /rpt_family="B2"
14411..14550
repeat_region /rpt_family="Alu"
17393..17539
repeat_region /rpt_family="Alu"
17784..17977
repeat_region /rpt_family="B2"
18370..18480
repeat_region /rpt_family="B4"
18485..18614
repeat_region /rpt_family="Alu"
19195..19364
repeat_region /rpt_family="B2"
19201..19273
trna /product="tRNA-Ser"
Sc=11.21"
19364..20105
repeat_region /rpt_family="Alu"
20474..20688
repeat_region /rpt_family="B2"
20770..20868
repeat_region /rpt_family="Alu"
20960..21086
repeat_region /rpt_family="MIR"
21223..21415
repeat_region /rpt_family="B2"
21600..21746
repeat_region /rpt_family="Alu"
21991..22060
repeat_region /rpt_family="ID"
22845..22953
repeat_region /rpt_family="Alu"
22965..23123
repeat_region /rpt_family="B2"
23145..23338
repeat_region /rpt_family="B2"
23760..23895
repeat_region /rpt_family="B2"
23899..24191
repeat_region /rpt_family="MaLR"
24476..24585
repeat_region /rpt_family="B2"
24672..24818
repeat_region /rpt_family="Alu"
25228..25408
repeat_region /rpt_family="B2"
25503..25688
repeat_region /rpt_family="B2"
25796..25880
repeat_region /rpt_family="Alu"
25986..26135
repeat_region /rpt_family="B4"
26231..26740
repeat_region /rpt_family="L1"
26750..26900
repeat_region /rpt_family="Alu"
29766..29872
repeat_region /rpt_family="Alu"
30031..30314
repeat_region /rpt_family="B4"
31450..31578
repeat_region /rpt_family="Alu"
32188..32267
repeat_region /rpt_family="ID"
32299..32426
repeat_region /rpt_family="B4"
32479..32676
repeat_region /rpt_family="B4"
32909..33004
repeat_region /rpt_family="L1"
19364..20105
20474..20688
20770..20868
20960..21086
21223..21415
21600..21746
21991..22060
22845..22953
22965..23123
23145..23338
23760..23895
23899..24191
24476..24585
24672..24818
25228..25408
25503..25688
25796..25880
25986..26135
26231..26740
26750..26900
29766..29872
30031..30314
31450..31578
32188..32267
32299..32426
32479..32676
32909..33004
```

```
repeat_region 33598..33685
Query Match 1.3%; Score 42; DB 6; Length 162528;
Best Local Similarity 100.0%; Pred. No. 9.1e-12;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1999 AAAACAAGGAATTTTGGCCACCATTTTGGATTTCAGCAGGT 2040
|||||
Db 118998 AAAACAAGGAATTTTGGCCACCATTTTGGATTTCAGCAGGT 119039
|||||

RESULT 75
AR372392/c AR372392 612 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 31 from patent US 6395889.
ACCESSION AR372392
VERSION AR372392.1 GI:34609704
KEYWORDS SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 612)
TITLES Robison,K.E.
JOURNAL Nucleic acid molecules encoding human protease homologs
Patent: US 6395889-A 31 28-MAY-2002;
Millennium Pharmaceuticals, Inc.; Cambridge, MA
FEATURES Location/Qualifiers
source
1..612
/organism="unknown"
/mol_type="genomic DNA"

Query Match 1.1%; Score 34; DB 2; Length 612;
Best Local Similarity 100.0%; Pred. No. 4.2e-07;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2495 CTGGGGCCCCAGTCACCTGTGTGATCTTCTATGA 2528
|||||
Db 365 CTGGGGCCCCAGTCACCTGTGTGATCTTCTATGA 332
|||||

RESULT 76
AC122776 72149 bp DNA linear HTG 25-MAY-2002
DEFINITION Mus musculus clone RP24-201E8, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC122776
VERSION AC122776.1 GI:21206389
KEYWORDS HTG; HTGS PHASE0.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 72149)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLES Mus musculus, clone RP24-201E8
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 72149)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Charaz,B., Choepel,F., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D.,
Galecki,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazar,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Minova,T., Menga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicoli,K., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
```


TITLE
JOURNAL
COMMENT

Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (25-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L26573
Center clone name: 201_E_8

* NOTE: This record contains 88 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 733: contig of 733 bp in length
833: gap of 100 bp
834 1551: contig of 718 bp in length
1552 1651: gap of 100 bp
1652 2370: contig of 719 bp in length
2371 2470: gap of 100 bp
2471 3207: contig of 737 bp in length
3208 3307: gap of 100 bp
3308 4023: contig of 716 bp in length
4024 4123: gap of 100 bp
4124 4833: contig of 710 bp in length
4834 4933: gap of 100 bp
4934 5636: contig of 703 bp in length
5637 5736: gap of 100 bp
5737 6411: contig of 675 bp in length
6412 6511: gap of 100 bp
6512 7248: contig of 737 bp in length
7249 7348: gap of 100 bp
7349 8088: contig of 740 bp in length
8089 8188: gap of 100 bp
8189 8919: contig of 731 bp in length
8920 9019: gap of 100 bp
9020 9758: contig of 739 bp in length
9759 9858: gap of 100 bp
9859 10588: contig of 730 bp in length
10589 10688: gap of 100 bp
10689 11420: contig of 732 bp in length
11421 11520: gap of 100 bp
11521 12236: contig of 716 bp in length
12237 12336: gap of 100 bp
12337 13033: contig of 697 bp in length
13034 13133: gap of 100 bp
13134 13866: contig of 733 bp in length
13867 13966: gap of 100 bp
13967 14670: contig of 704 bp in length
14671 14770: gap of 100 bp
14771 15475: contig of 705 bp in length
15476 15575: gap of 100 bp
15576 16289: contig of 714 bp in length
16289 16389: gap of 100 bp
16390 16290

16390 17130: contig of 741 bp in length
17131 17230: gap of 100 bp
17231 17963: contig of 733 bp in length
17964 18063: gap of 100 bp
18064 18768: contig of 705 bp in length
18769 18868: gap of 100 bp
18869 19597: contig of 729 bp in length
19598 19697: gap of 100 bp
19698 20447: contig of 750 bp in length
20448 20547: gap of 100 bp
20549 21283: contig of 736 bp in length
21284 21383: gap of 100 bp
21384 22121: contig of 738 bp in length
22122 22221: gap of 100 bp
22222 22943: contig of 722 bp in length
22944 23043: gap of 100 bp
23044 23755: contig of 712 bp in length
23756 23855: gap of 100 bp
23856 24596: contig of 741 bp in length
24597 24696: gap of 100 bp
24697 25409: contig of 713 bp in length
25410 25509: gap of 100 bp
25510 26208: contig of 699 bp in length
26209 26308: gap of 100 bp
26309 27001: contig of 693 bp in length
27002 27101: gap of 100 bp
27102 27831: contig of 730 bp in length
27832 27931: gap of 100 bp
27932 28661: contig of 730 bp in length
28662 28761: gap of 100 bp
28762 29500: contig of 739 bp in length
29501 29600: gap of 100 bp
29601 30333: contig of 733 bp in length
30334 30433: gap of 100 bp
30434 31135: contig of 702 bp in length
31136 31235: gap of 100 bp
31236 31945: contig of 710 bp in length
31946 32045: gap of 100 bp
32046 32757: contig of 712 bp in length
32758 32857: gap of 100 bp
32858 33581: contig of 724 bp in length
33582 33681: gap of 100 bp
33682 34361: contig of 680 bp in length
34362 34461: gap of 100 bp
34462 35166: contig of 705 bp in length
35167 35266: gap of 100 bp
35267 35993: contig of 727 bp in length
35994 36093: gap of 100 bp
36094 36829: contig of 736 bp in length
36830 36929: gap of 100 bp
36930 37662: contig of 733 bp in length
37663 37762: gap of 100 bp
37763 38501: contig of 739 bp in length
38502 38601: gap of 100 bp
38602 39342: contig of 741 bp in length
39343 39442: gap of 100 bp
39443 40161: contig of 719 bp in length
40162 40261: gap of 100 bp
40262 40982: contig of 721 bp in length
40983 41082: gap of 100 bp
41083 41816: contig of 734 bp in length
41817 41916: gap of 100 bp
41917 42623: contig of 707 bp in length
42624 42723: gap of 100 bp
42724 43429: contig of 706 bp in length
43430 43529: gap of 100 bp
43530 44262: contig of 733 bp in length
44263 44362: gap of 100 bp
44363 45097: contig of 735 bp in length
45098 45197: gap of 100 bp
45198 45926: contig of 729 bp in length
45927 46026: gap of 100 bp
46027 46744: contig of 718 bp in length
46744 46027

```

* 46745 46844: gap of 100 bp
* 47580: contig of 736 bp in length
* 47680: gap of 100 bp
* 47681: contig of 738 bp in length
* 48418: contig of 706 bp in length
* 48519: gap of 100 bp
* 49224: contig of 706 bp in length
* 49225: gap of 100 bp
* 49325: contig of 718 bp in length
* 50043: gap of 100 bp
* 50143: contig of 734 bp in length
* 50877: gap of 100 bp
* 50977: gap of 100 bp
* 51693: contig of 716 bp in length
* 51793: contig of 704 bp in length
* 52497: gap of 100 bp
* 52597: contig of 710 bp in length
* 53307: gap of 100 bp
* 53407: contig of 704 bp in length
* 54111: gap of 100 bp
* 54211: contig of 738 bp in length
* 54949: gap of 100 bp
* 55049: contig of 728 bp in length
* 55777: gap of 100 bp
* 55877 56620: contig of 744 bp in length

Query Match
Best Local Similarity 1.0%; Score 33; DB 12; Length 72149;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2008 GAATTTGGGCCACCATTTTGGATTGAGCAGGT 2040
Db 1714 GAATTTGGGCCACCATTTTGGATTGAGCAGGT 1746

RESULT 77
LOCUS CR353491 1449 bp mRNA linear VRT 11-MAR-2004
DEFINITION Gallus gallus finished cDNA, clone CHEST24919.
ACCESSION CR353491
VERSION CR353491.1 GI:45424761
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus
1 (bases 1 to 1449)
Boardman, P.E., Bonfield, J.K., Brown, W.R.A., Carder, C., Chalk, S.E.,
Croning, M.D.R., Davies, R.M., Francie, M.D., Grafham, D.V.,
Hubbard, S.D., Humphray, S.J., Hunt, P.J., Maddison, M., McLaren, S.R.,
Niblett, D., Paterson, I.M., Rogers, J., Scott, C.E., Taylor, R.G.,
Tickle, C., and Wilson, S.A.
Direct Submission
Submitted (09-MAR-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chicken@bms.umist.ac.uk
BBSRC/BuNdee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
sequencing project.
This sequence is from the
BBSRC/BuNdee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from liver, normalised, and poly A-trimmed.
EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
coli DH10B.
Location/Qualifiers
1..1449
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST24919"
/clone_lib="CSEQHN33"
/dev_stage="adult"
FEATURES
source

```

ORIGIN

```

Query Match
Best Local Similarity 1.0%; Score 32; DB 11; Length 1449;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2458 TTAATGCAGAGTCAGATATCTTCAGGGTTCC 2489
Db 815 TTAATGCAGAGTCAGATATCTTCAGGGTTCC 846

```

RESULT 78

```

AC163387 194188 bp DNA linear ROD 13-JUL-2005
LOCUS Mus musculus chromosome 1, clone RP24-181G14, complete sequence.
DEFINITION AC163387
ACCESSION AC163387
VERSION AC163387.2 GI:70778622
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 194188)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 1, clone RP24-181G14
Unpublished
2 (bases 1 to 194188)

```

REFERENCE

```

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Atchchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J.,
Choepe, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D.,
Galgani, J., Gardyna, S., Graham, L., Horton, L., Hulme, W., Hafez, N.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kell, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
McCarthy, M., McDonald, P., Major, J., Manning, J., Matthews, C.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (10-JUN-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 194188)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Atchchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J.,
Choepe, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D.,
Galgani, J., Gardyna, S., Graham, L., Horton, L., Hulme, W., Hafez, N.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kell, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
McCarthy, M., McDonald, P., Major, J., Manning, J., Matthews, C.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (10-JUN-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 194188)

```

TITLE

JOURNAL

REFERENCE

AUTHORS

		repeat_region	/rpt_family="RLTR1"		
		repeat_region	37290..37986		
		repeat_region	/rpt_family="Lx4"		
		repeat_region	37997..38021		
		repeat_region	/rpt_family="(TAAA)n"		
		repeat_region	38217..38258		
		repeat_region	/rpt_family="(TG)n"		
		repeat_region	38358..38318		
		repeat_region	/rpt_family="(GA)n"		
		repeat_region	38559..39775		
		repeat_region	/rpt_family="Lx4"		
		repeat_region	39792..39892		
		repeat_region	/rpt_family="RLTR13A"		
		repeat_region	39893..40769		
		repeat_region	/rpt_family="RLTR13B"		
		repeat_region	40773..41089		
		repeat_region	/rpt_family="Lx4"		
		repeat_region	41101..41169		
		repeat_region	/rpt_family="CT-rich"		
		repeat_region	complement(41175..41247)		
		repeat_region	/rpt_family="RLTR11A"		
		repeat_region	41248..41284		
		repeat_region	/rpt_family="(CA)n"		
		repeat_region	complement(41285..41570)		
		repeat_region	/rpt_family="RLTR11A"		
		repeat_region	41581..42364		
		repeat_region	/rpt_family="LiVL2"		
		repeat_region	42364..42418		
		repeat_region	/rpt_family="L1"		
		repeat_region	42841..44071		
		repeat_region	/rpt_family="L1"		
		repeat_region	44072..44115		
		repeat_region	/rpt_family="(TAAA)n"		
		repeat_region	44116..46091		
		repeat_region	/rpt_family="L1"		
		repeat_region	complement(46092..46283)		
		repeat_region	/rpt_family="MT2C"		
		repeat_region	46876..47091		
		repeat_region	/rpt_family="B3"		
		repeat_region	47134..47324		
		repeat_region	/rpt_family="(TCTCCC)n"		
		repeat_region	48006..48153		
		repeat_region	/rpt_family="Bi_MM"		
		repeat_region	48154..48208		
		repeat_region	/rpt_family="(CAAAAA)n"		
		repeat_region	complement(48286..48496)		
		repeat_region	/rpt_family="RMER3"		
		repeat_region	49024..49100		
		repeat_region	/rpt_family="(TC)n"		
				Query Match	0.9%; Score 29; DB 6
				Best Local Similarity	100.0%; Pred. NO. 0.0%
				Matches	29; Conservative 0; Mismatches
					1072 TATGAGAAATGATGATCTGAGGTGGA 1100


```

methods of use thereof
JOURNAL Patent: WO 0147944-A 971 05-JUL-2001;
FEATURES Curagen Corporation (US)
source Location/Qualifiers
1. .51
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Accession number cg43920127"

ORIGIN
Query Match 0.8%; Score 25; DB 2; Length 51;
Best Local Similarity 100.0%; Pred.No. 0.082;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2756 GACACAGCATAAGAGTTCCTGAATC 2780
Db 51 GACACAGCATAAGAGTTCCTGAATC 27

RESULT 82
LOCUS CQ572930 9154 bp DNA linear PAT 02-FEB-2004
DEFINITION Sequence 688 from Patent WO0171042.
ACCESSION CQ572930
VERSION CQ572930.1 GI:41637266
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.
TITLE Detection kits, such as nucleic acid arrays, for detecting the
expression of 10,000 or more Drosophila genes and uses thereof
JOURNAL Patent: WO 0171042-A 688 27-SEP-2001;
PE Corporation (NY) (US)
FEATURES
source
1. .9154
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"

ORIGIN
Query Match 0.8%; Score 24; DB 2; Length 9154;
Best Local Similarity 100.0%; Pred.No. 0.36;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 625 CTATTAGAGAGAAAGAAACGCATT 648
Db 1771 CTATTAGAGAGAAAGAAACGCATT 1794

RESULT 83
LOCUS CQ608996 40388 bp DNA linear PAT 02-FEB-2004
DEFINITION Sequence 36754 from Patent WO0171042.
ACCESSION CQ608996
VERSION CQ608996.1 GI:41661310
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.
TITLE Detection kits, such as nucleic acid arrays, for detecting the
expression of 10,000 or more Drosophila genes and uses thereof
JOURNAL Patent: WO 0171042-A 36754 27-SEP-2001;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers

methods of use thereof
JOURNAL Patent: WO 0147944-A 971 05-JUL-2001;
FEATURES Curagen Corporation (US)
source Location/Qualifiers
1. .51
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Accession number cg43920127"

ORIGIN
Query Match 0.8%; Score 24; DB 2; Length 40388;
Best Local Similarity 100.0%; Pred.No. 0.37;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 625 CTATTAGAGAGAAAGAAACGCATT 648
Db 13418 CTATTAGAGAGAAAGAAACGCATT 13441

RESULT 84
LOCUS AC015386 79001 bp DNA linear HTG 16-NOV-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC015386
VERSION AC015386.1 GI:6435949
KEYWORDS HTG; HTGS PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 79001)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDM:10210324 by the submitter.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. .79001
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

ORIGIN
Query Match 0.8%; Score 24; DB 12; Length 79001;
Best Local Similarity 100.0%; Pred.No. 0.37;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 625 CTATTAGAGAGAAAGAAACGCATT 648
Db 68006 CTATTAGAGAGAAAGAAACGCATT 68029

RESULT 85
LOCUS AC090436 102653 bp DNA linear PLN 01-OCT-2002
DEFINITION Chlamydomonas reinhardtii clone cr-3hi, complete sequence.
ACCESSION AC090436
VERSION AC090436.37 GI:23396228
KEYWORDS HTG.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 102653)
AUTHORS Jia, H., Lin, S., Wu, H., Dutcher, S. and Roe, B.A.
TITLE Chlamydomonas reinhardtii BAC Clone cr-3hi
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 102653)
AUTHORS Jia, H., Lin, S., Wu, H., Dutcher, S. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2001) Department Of Chemistry And Biochemistry,

```

REFERENCE
AUTHORS
TITLE
JOURNAL

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 102653)
Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (05-SEP-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 102653)
Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (05-SEP-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 102653)
Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (01-OCT-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Oct 1, 2002 this sequence version replaced gi:22748433.

Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

COMMENT

FEATURES
source
Location/Qualifiers
1..102653
/organism="Chlamydomonas reinhardtii"
/mol_type="genomic DNA"
/db_xref="taxon:3055"
/clone="cr-3h1"
/clone_lib="cr-3h1"

ORIGIN

Query Match 0.8%; Score 24; DB 4; Length 102653;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 GCCGCTGCTACTGCCCGCTGCT 49
|||||
Db 18109 GCCGCTGCTACTGCCCGCTGCT 18086
|||||

RESULT 86
AC090434/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC090434 127686 bp DNA linear PLN 17-SEP-2003
Chlamydomonas reinhardtii clone cr-1j6, complete sequence.
AC090434.43 GI:34787433
HTG.
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
Chlamydomonadales; Chlamydomonadales; Chlamydomonadales.

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 127686)
Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
Chlamydomonas reinhardtii BAC Clone cr-1j6
Unpublished

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 127686)
Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (21-FEB-2001) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 127686)
Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (03-AUG-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS

4 (bases 1 to 127686)
Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.

TITLE
JOURNAL

Direct Submission
Submitted (14-AUG-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

5 (bases 1 to 127686)
Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (17-SEP-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT

On Sep 17, 2003 this sequence version replaced gi:22218469.

Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

FEATURES
source

Location/Qualifiers
1..127686
/organism="Chlamydomonas reinhardtii"
/mol_type="genomic DNA"
/db_xref="taxon:3055"
/clone="cr-1j6"
/clone_lib="cr-1j6"
/note="This is one of two clone in the same well cr-1j6"

ORIGIN

Query Match 0.8%; Score 24; DB 4; Length 127686;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 GCCGCTGCTACTGCCCGCTGCT 49
|||||
Db 102737 GCCGCTGCTACTGCCCGCTGCT 102714
|||||

RESULT 87
AC010110
LOCUS

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC010110 172372 bp DNA linear INV 24-JUL-2002
Drosophila melanogaster 3L BAC RP98-9N2 (Roswell Park Cancer
Institute Drosophila BAC Library) complete sequence.
AC010110
AC010110.7 GI:21553183
HTG.
Drosophila melanogaster (fruit fly)

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS

1 (bases 1 to 172372)
Muzny,D., Scherer,S., Adams,M.D., Holt,R.A., Evans,C.A.,
Gocayne,J.D., Tabor,P., Williamson,A., Homsai,F.H., Chen,C.,
Dugan-Rocha,S.D., Sodergren,E.S., Hodgson,A.H., Chen,R.C.,
Ayele,M., Scott,G.S., Worley,K.W., Anamastides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Beeson,K.Y., Brown,M., Bunay,C.,
Busam,D.A., Center,A., Chen,G., Chen,Z., Clerc-Blankenburg,K.,
Davenport,L.B., Dietz,S.M., Ding,Y., Dodson,K., Doup,L.E.,
Draper,H., Emery-Cohen,A., Fierster,S., Garg,N.D.S., Houck,J.,
Hostin,D., Howland,T.J., Hume,J., Ibegwam,C., Jaisi,M., Kovar,C.,
Liu,W., Mattei,B., McIntosh,R.C., Morgan,M., Moy,M., Murphy,B.,
Nelson,K.A., Ndasya,Y., Nguyen,N., Perez,L., Pittman,G.S., Puri,V.,
Scheeler,F., Shih,H., Strong,R., Tector,C., Wang,Q., Williams,S.M.,
Xiang,J., Zavitz,J.S., Zhou,J., Zorrilla,S., Smith,H.O.,
Wheeler,D., Weinstock,G., Gibbs,R. and Venter,J.C.
Direct Submission

TITLE
JOURNAL

Unpublished
2 (bases 1 to 172372)
Worley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amarutunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowles,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Cartron,T.F.,
Cartier,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.ncbi.nlm.nih.gov/genomes/info/genbank/annotation.html>

REFERENCE AUTHORS

ORIGIN

```

Query Match      0.8%; Score 24; DB 13; Length 172372;
Best Local Similarity 100.0%; Pred. NO. 0.38;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 625 CTATTACAGAAAGAAACGCATT 648
Db 88694 CTATTACAGAAAGAAACGCATT 89717

RESULT 88
AC127924
LOCUS      298235 bp DNA linear HTG 19-NOV-2002
DEFINITION Rattus norvegicus clone CH230-23318, *** SEQUENCING IN PROGRESS
AC127924
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
VERSION    Rattus norvegicus (Norway rat)
KEYWORDS   Rattus norvegicus
SOURCE     Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
           Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
           1 (bases 1 to 298235)
Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Amin,A., Anguiano,D., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Ayodeji,M., Baca,E., Baden,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Barber,M., Barneshead,M., Benahmed,F.,
Biswal,D., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,J., Cesar,H., Chen,R., Chen,Z.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,X., Dinh,H., Divya,K.,
Draher,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebrgeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,M., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlik,P., Hawes,A., Henderson,N., Hernandez,J.,
Henderson,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangun,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokeme,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfanukoeh,C.,
Plummer,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puzo,M., Quiroz,J., Rachin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,P.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Shen,H.,
Sander,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soza,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
Valas,R., Vexa,V., Villanasa,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Unpublished
2 (bases 1 to 298235)
Worley,K.C.

Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 298235)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23915349.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

Center: Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: KAAS
Center clone name: CH230-23318

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 245203 bases at least Q40
Consensus quality: 249375 bases at least Q30
Consensus quality: 252157 bases at least Q20
Estimated insert size: 248526; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases, sum-of-contigs estimation

***** NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

***** NOTE: contig of 9520 bp in length
* 9521
* 9520: gap of unknown length
* 9521: contig of 15345 bp in length
* 24966
* 25065: gap of unknown length
* 25066
* 43623: contig of 18558 bp in length
* 43624
* 43723: gap of unknown length
* 43724
* 58040: contig of 14317 bp in length
* 58041
* 58140: gap of unknown length
* 58141
* 276219: contig of 218079 bp in length
* 276220
* 276319: gap of unknown length
* 276320
* 281271: contig of 4952 bp in length
* 281272
* 281371: gap of unknown length
* 281372
* 282614: contig of 1243 bp in length
* 282615
* 282714: gap of unknown length
* 282715
* 283928: contig of 1214 bp in length
* 283929
* 284028: gap of unknown length
* 284029
* 285169: contig of 1141 bp in length
* 285170
* 285269: gap of unknown length
* 285270
* 286516: contig of 1247 bp in length
* 286517
* 286516: gap of unknown length

* 286617 287814: contig of 1198 bp in length
 * 287915 287914: gap of unknown length
 * 287915 289338: contig of 1024 bp in length
 * 289339 289338: gap of unknown length
 * 289339 290451: contig of 1413 bp in length
 * 290452 290551: gap of unknown length
 * 290552 292527: contig of 1976 bp in length
 * 292528 292627: gap of unknown length
 * 292628 298235: contig of 5608 bp in length.

FEATURES

Location/Qualifiers
 source
 1. 298235
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-23318"

misc_feature

1. 1277
 /note="wgs end extension
 clone_end:Sp6"

misc_feature

complement(2544..3428)
 /note="clone_boundary
 clone_end:Sp6
 site:
 end sequence:192091135"

gap

9521..9620
 /estimated_length=unknown
 complement(11784..12687)

misc_feature

clone_end:T7
 end sequence:192091133"
 24966..25065
 /estimated_length=unknown
 43624..43723
 /estimated_length=unknown
 58041..58140
 /estimated_length=unknown
 276220..276319
 /estimated_length=unknown
 281372..281371
 /estimated_length=unknown
 282615..282714
 /estimated_length=unknown
 283929..284028
 /estimated_length=unknown
 285170..285269
 /estimated_length=unknown
 286517..286616
 /estimated_length=unknown
 287815..287914
 /estimated_length=unknown
 288939..289038
 /estimated_length=unknown
 290452..290551
 /estimated_length=unknown
 292528..292627
 /estimated_length=unknown

ORIGIN

Query Match 0.8%; Score 24; DB 12; Length 298235;
 Best Local Similarity 100.0%; Pred. No. 0.39;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 GCCGCTGCTACTGCGCGCTGCT 49
 |||||

Db 221166 GCCGCTGCTACTGCGCGCTGCT 221189

RESULT 89

AE003475

LOCUS

DEFINITION

ACCESSION

301691 bp DNA linear INV 13-JAN-2006
 Drosophila melanogaster chromosome 3L, section 9 of 83 of the
 complete sequence.
 AE003475 AE002584 AF014296

AE003475.4 GI:55380495

VERSION
KEYWORDS

Source Drosophila melanogaster (fruit fly)

ORGANISM

Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

1 (bases 1 to 301691)
 Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
 Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
 George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Chen,L.X.,
 Sutton,G.G., Wortman,J.R., Vandal,M.D., Zhang,Q., Chen,L.X.,
 Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
 Wan,K.H., Doyle,C., Baxter,E.G., Heit,G., Nelson,C.R., Gabor,G.L.,
 Abail,J.F., Agbayani,A., An,H.J., Andrews-Frankoch,C., Baldwin,D.,
 Balow,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M.,
 Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S.,
 Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P.,
 Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A.,
 Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B.,
 Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I.,
 Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S.,
 Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C.,
 Ferrieri,S., Fleischmann,W., Fosler,C., Gabriellian,A.E., Garg,N.S.,
 Gelbart,W.M., Glasser,K., Glodek,A., Gong,E., Gorrell,J.H., Gu,Z.,
 Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
 Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
 Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
 Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
 Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
 Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
 McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
 Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
 Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
 Nuskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
 Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
 Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamoe,I.,
 Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
 Stapleton,M., Strong,R., Sun,E., Svirskaas,R., Tector,C., Turner,R.,
 Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
 Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
 Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
 Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
 Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
 Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
 The genome sequence of Drosophila melanogaster
 Science 287 (5461), 2185-2195 (2000)

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

2 (bases 1 to 301691)
 Celniker,S.E., Wheeler,D.A., Krommiller,B., Carlson,J.W.,
 Halpern,A., Patel,S., Adams,M., Champe,M., Dugan,S.P., Friese,E.,
 Hodgson,A., George,R.A., Hoskins,R.A., Laverty,T., Muzny,D.M.,
 Nelson,C.R., Pacleb,J.M., Park,S., Pfeiffer,B.D., Richards,S.,
 Sodergren,E.J., Svirskaas,R., Tabor,P.E., Wan,K., Stapleton,M.,
 Sutton,G.G., Venter,C., Weinstock,G., Scherer,S.E., Myers,E.W.,
 Gibbs,R.A. and Rubin,G.M.
 Finishing a whole-genome shotgun: release 3 of the Drosophila
 melanogaster euchromatic genome sequence
 Genome Biol. 3 (12), RESEARCH0079 (2002)
 12537568

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

3 (bases 1 to 301691)
 Misra,S., Crosby,M.A., Mungall,C.J., Matthews,B.B., Campbell,K.S.,
 Hradecky,P., Huang,Y., Kaminker,J.S., Millburn,G.H., Prochnik,S.E.,
 Smith,C.D., Tupy,J.L., Whitfield,E.J., Bayraktaroglu,L.,
 Berman,B.P., Bettencourt,B.R., Celniker,S.E., de Grey,A.D.,
 Drysdale,R.A., Harris,N.L., Richter,J., Russo,S., Schroeder,A.J.,
 Shu,S.Q., Stapleton,M., Yamada,C., Ashburner,M., Gelbart,W.M.,
 Rubin,G.M. and Lewis,S.E.
 Annotation of the Drosophila melanogaster euchromatic genome: a
 systematic review
 Genome Biol. 3 (12), RESEARCH0083 (2002)
 12537572

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

4 (bases 1 to 301691)

AUTHORS	Kaminker, J.S., Bergman, C.M., Kronmiller, B., Carlson, J., Svirskaas, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E., Rubin, G.M., Ashburner, M. and Celnikier, S.:		
	The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective		
JOURNAL	Genome Biol.	3 (12), RESEARCH0084	(2002)
PUBMED	12537573		
REFERENCE	5	(bases 1 to 301691)	
AUTHORS	Celnikier, S., Carlson, J., Wan, K., Pfeiffer, B., Frise, E., George, R., Hoskins, R., Stapleton, M., Pacleb, J., Park, S., Svirskaas, R., Smith, E., Yu, C. and Rubin, G.		
	Berkeley Drosophila Genome Project		
CONSRSTM	Drosophila melanogaster release 4 sequence		
JOURNAL	Unpublished		
REFERENCE	6	(bases 1 to 301691)	
AUTHORS	Adams, M.D., Celnikier, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.		
	Direct Submissions		
JOURNAL	Submitted (21-MAR-2000)	Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA	
REFERENCE	7	(bases 1 to 301691)	
AUTHORS	FlyBase		
	Direct Submission		
JOURNAL	Submitted (05-JAN-2006)	FlyBase, Harvard University, Biological Laboratories, 16 Divinity Ave, Cambridge, MA 02138, USA	
COMMENT	On Nov 4, 2004 this sequence version replaced gi:23092840.		
FEATURES	Location/Qualifiers		
	1..301691		
source	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
gene	/db_xref="taxon:7227"		
	<267..>1262		
gene	/gene="CG33231"		
	/locus_tag="Dmel_CG33231"		
mRNA	/old_locus_tag="CG33231"		
	/map="62E3-62E3"		
mRNA	/db_xref="FLYBASE:FBgn0053231"		
	join(<267..688,1139..>1262)		
CDS	/gene="CG33231"		
	/locus_tag="Dmel_CG33231"		
CDS	/old_locus_tag="CG33231"		
	/product="CG33231-RA"		
CDS	/db_xref="FLYBASE:FBgn0053231"		
	join(267..688,1139..>1262)		
CDS	/gene="CG33231"		
	/locus_tag="Dmel_CG33231"		
CDS	/old_locus_tag="CG33231"		
	/note="CG33231 gene product from transcript CG33231-RA"		
CDS	/codon_start=1		
	/product="CG33231-PA"		
CDS	/protein_id="AA564938.1"		
	/db_xref="GI:45445754"		
CDS	/db_xref="FLYBASE:FBgn0053231"		
	/translation="MGMGRGIMVLDLDELKRLKRSHTATVSNLRKSGATSDATRAA		
CDS	SDEVNQRNLQILIKNSVKNSTFKHDDAVSLVRLNLSTLQGRSLARDDAAGCAGA		
	SEGSSSGRETEAIIKNSAVARRRQQQQQQQQQDPGGHNSIRIRPSNKANSSSS		
CDS	QCCSCSNRPVTSFSPDIL"		
	2014..34339		
gene	/gene="CG33232"		
	/locus_tag="Dmel_CG33232"		
gene	/old_locus_tag="CG33232"		
	/map="62E3-62E4"		
mRNA	/db_xref="FLYBASE:FBgn0053347"		
	join(2014..4380,19675..19948,20184..20339,30122..30263,		
mRNA	30320..30840,30912..31069,31129..33147,33210..33457,		
	33517..34339)		
mRNA	/gene="CG33232"		
	/locus_tag="Dmel_CG33232"		
mRNA	/old_locus_tag="CG33232"		
	/product="CG33232-RD, transcript variant D"		
gene	/db_xref="FLYBASE:FBgn0053347"		
	complement(<9878..>10705)		
mRNA	/gene="CG16762"		
	/locus_tag="Dmel_CG16762"		
mRNA	/old_locus_tag="CG16762"		
	/map="62E3-62E3"		
mRNA	/db_xref="FLYBASE:FBgn0035343"		
	complement(join(<9878..10228,10292..>10705))		
mRNA	/gene="CG16762"		
	/locus_tag="Dmel_CG16762"		
mRNA	/old_locus_tag="CG16762"		
	/product="CG16762-PA"		
mRNA	/db_xref="FLYBASE:FBgn0035343"		
	complement(join(9878..10228,10292..10705))		
mRNA	/gene="CG16762"		
	/locus_tag="Dmel_CG16762"		
mRNA	/old_locus_tag="CG16762"		
	/note="CG16762 gene product from transcript CG16762-RA"		
mRNA	/codon_start=1		
	/product="CG16762-PA"		
mRNA	/protein_id="AAF47648.1"		
	/db_xref="GI:7292239"		
mRNA	/db_xref="FLYBASE:FBgn0035343"		
	/translation="MCEMSVKILASLAMIQLLVLSPAEGISRLPPVVGILPTODGL		
mRNA	TRLPLNSRVTDTPFASVACFGYIGESNLIAELYSANVTKCYNAADSRKGDADFL		
	ATRTILRSERVCSBLRACNELNTLTESFQCHAVNGSNTVSTYISGNASESASVL		
mRNA	EERYRVDLRHGCCQCORARHYVVESTARNYVLOACLDRAPKPKMPKPTSTTSSTT		
	TTTTTTTTEAPPTPTTTEAPLNVEDQFKQLNLLN"		
mRNA	complement(11253..13604)		
	/gene="Cyp4d20"		
mRNA	/locus_tag="Dmel_CG16761"		
	/old_locus_tag="CG16761"		
mRNA	/map="62E3-62E3"		
	/db_xref="FLYBASE:FBgn0035344"		
mRNA	complement(join(11253..12229,12296..12492,12555..12624,		
	12801..13187,13257..13604))		
mRNA	/gene="Cyp4d20"		
	/locus_tag="Dmel_CG16761"		
mRNA	/old_locus_tag="CG16761"		
	/product="CG16761-RA"		
mRNA	/db_xref="FLYBASE:FBgn0035344"		
	complement(join(11514..12229,12296..12492,12555..12624,		
mRNA	12801..13187,13257..13419))		
	/gene="Cyp4d20"		
mRNA	/locus_tag="Dmel_CG16761"		
	/old_locus_tag="CG16761"		
mRNA	/note="CG16761 gene product from transcript CG16761-RA"		
	/codon_start=1		
mRNA	/product="CG16761-PA"		
	/protein_id="RAP47649.1"		
mRNA	/db_xref="GI:72922409"		
	/db_xref="FLYBASE:FBgn0035344"		
mRNA	Query Match 0.8%; Score 24; DB 13; Length 301691;		
	Best Local Similarity 100.0%; Pred.No. 0.39;		
mRNA	Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
	QY 625 CTATTAAAGAGAAAGAACGCATT 648		
mRNA	Db 275631 CTATTAAAGAGAAAGAACGCATT 275654		
	RESULT 90		
mRNA	CS196104		
	LOCUS		
mRNA	Sequence 3 from Patent WO2005106021.		
	Accession CS196104		
mRNA	Version CS196104.1		
	GI:83410069		
mRNA	Keywords		
	synthetic construct		
mRNA	Other sequences; artificial sequences.		
	REFERENCE		
mRNA	AUTHORS		
	JOURNAL		
mRNA	Patent: WO 2005106021-A 3 10-NOV-2005;		

Thu Jun 22 09:04:25 2006

```

FEATURES
  source
    Bayer HealthCare AG (DE)
    Location/Qualifiers
      1..23
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /note="forward primer"

ORIGIN
  Query Match
  Best Local Similarity 100.0%; Score 23; DB 2; Length 23;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 887 TCGTAACACAGAGAAGAGGAGA 909
Db 1 TCGTAACACAGAGAAGAGGAGA 23

RESULT 91
LOCUS AR631300 29 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 45 from patent US 6844180.
ACCESSION AR631300
VERSION AR631300.1 GI:59770958
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
  1 (bases 1 to 29)
  Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
  Serine protease genes related to DPPIV
  Patent: US 6844180-A 45 18-JAN-2005;
  Ferring BV;
  NLX;
FEATURES
  source
    Location/Qualifiers
      1..29
      /organism="unknown"
      /mol_type="genomic DNA"

ORIGIN
  Query Match
  Best Local Similarity 100.0%; Score 23; DB 2; Length 29;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 CATGGCAGCAGCAATGGAAACAG 235
Db 7 CATGGCAGCAGCAATGGAAACAG 29

RESULT 92
LOCUS AX608768 29 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 45 from Patent WO0231134.
ACCESSION AX608768
VERSION AX608768.1 GI:28404320
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
  1
  Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.-L.
  Novel serine protease genes related to dppiv
  Patent: WO 0231134-A 45 18-APR-2002;
  Ferring BV (NL)
FEATURES
  source
    Location/Qualifiers
      1..29
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"

ORIGIN
  Query Match
  Best Local Similarity 100.0%; Score 23; DB 2; Length 29;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 CATGGCAGCAGCAATGGAAACAG 235
Db 7 CATGGCAGCAGCAATGGAAACAG 29

RESULT 93
LOCUS BV055847 632 bp DNA linear STS 31-MAY-2003
DEFINITION S212P6100FP6.T0 CZECHII/Ei Mus musculus STS genomic, sequence
  tagged site.
ACCESSION BV055847
VERSION BV055847.1 GI:31171642
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
  1 (bases 1 to 632)
  Wade, C.M., Kulbokas, E.J. III, Kirby, A.W., Zody, M.C., Mullikin, J.C.,
  Lander, E.S., Lindblad-Toh, K. and Daly, M.J.
  The mosaic structure of variation in the laboratory mouse genome
  Nature 420 (6915), 574-578 (2002)
  12466852
CONTACT: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 632
Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGS CV3 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.
FEATURES
  source
    Location/Qualifiers
      1..632
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /strain="CZECHII/Ei"
      /db_xref="taxon:10090"
      /map="+" 1 21-583 7608303-7608864"
      /clone.lib="CZECHII/Ei"
      <1..>632

ORIGIN
  Query Match
  Best Local Similarity 100.0%; Score 23; DB 7; Length 632;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1477 ATCTATGAGAAACACACACAT 1499
Db 417 ATCTATGAGAAACACACACAT 439

RESULT 94
LOCUS BV311981 635 bp DNA linear STS 26-JAN-2005
DEFINITION S236P628FH12.T0 Alaskan Malamute Canis familiaris STS genomic,
  sequence tagged site.

```

ACCESSION BV311981
 VERSION BV311981.1 GI:57510473
 STS
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis
 1 (bases 1 to 635)
 The genome sequence of Canis familiaris
 Lindblad-Toh,K.
 Unpublished (2004)
 Contact: Kerstin Lindblad-Toh
 Whitehead Institute for Biomedical Research, Center for Genome Research
 320 Charles Street, Cambridge, MA 02141, USA
 Tel: 6172521477
 Fax: 6172580903
 Email: kersli@genome.wi.mit.edu
 Primer A: No sequence submitted
 Primer B: No sequence submitted
 STS size: 635
 Protocol:
 WGS-discovery (WGS):
 Paired-end and low-coverage whole genome shotgun reads were generated from 9 breeds
 (German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador Retriever, English Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese Water Dog -100,000 each)
 and five other canids (Chinese, Alaskan, Indian and Spanish Gray Wolf as well as the Californian Coyote).
 The WGS reads were placed uniquely on the CanFam1.0 boxer assembly and SNP detection was carried out by SSAHA-SNP. 863872 reads were annotated as STSs and 485941 SNPs were annotated with alleles from the boxer and the breed or canid from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.
 WGA-discovery (WGA) of Boxer/Poodle SNPs:
 A second set of SNPs was generated using a similar methodology except that the contigs from the 1.5x poodle assembly (Kirkness 2003) were used instead of WGS reads. Since this boxer assembly had quality scores, arbitrary quality scores of phred 40 were assigned before the poodle sequence was placed uniquely on the CanFam1.0 boxer assembly and SNP detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated with alleles from the boxer and the poodle. The validation rate for these SNPs was estimated at approximately 98%.
 A third set of SNPs were discovered by comparing reads in the WGA assembly. SNPs were defined as mismatch positions that had a base quality of >= 30 on both reads in a region that aligned without gaps, and with at most one additional mismatch in the ten flanking bases. For each allele, at least one additional read had to confirm it. 731476 SNPs were annotated with alleles between the two boxer alleles. The validation rate for these SNPs was estimated at approximately 98%.
 Location/Qualifiers
 1..635
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="AlaskanMalamute"
 /db_xref="taxon:9615"

ORIGIN
 STS
 Query Match 0.7%; Score 23; DB 7; Length 635;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3098 ATAAGANTTACTTAAAAA 3120
 Db 595 ATAGAGTACTTAAAAA 617
 RESULT 95
 CS211305/c
 LOCUS CS211305 1275 bp DNA linear PAT 15-DEC-2005
 DEFINITION Sequence 624 from Patent WO2005111246.
 ACCESSION CS211305
 VERSION CS211305.1 GI:83691295
 KEYWORDS Mesocricetus auratus (golden hamster)
 SOURCE Mesocricetus auratus
 ORGANISM Mesocricetus auratus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Cricetidae; Cricetinae; Mesocricetus.
 1
 Melville,M.W.
 Oligonucleotide arrays to monitor gene expression and methods for making and using same
 Patent: WO 2005111246-A 624 24-NOV-2005;
 Wyeth (US)
 FEATURES
 source
 1..1275
 /organism="Mesocricetus auratus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10036"
 misc_feature 389..407
 /note="n is a, c, g, or t"
 misc_feature 457..476
 /note="n is a, c, g, or t"
 ORIGIN
 Query Match 0.7%; Score 23; DB 2; Length 1275;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 26 GCCGCTGCTACTGCCGCGCTGC 48
 Db 501 GCCGCTGCTACTGCCGCGCTGC 479
 RESULT 96
 AB056577/c
 LOCUS AB056577 1275 bp mRNA linear ROD 25-SEP-2004
 DEFINITION Mesocricetus auratus mRNA for homeobox B3, complete cds.
 ACCESSION AB056577
 VERSION AB056577.1 GI:13359292
 KEYWORDS Mesocricetus auratus (golden hamster)
 SOURCE Mesocricetus auratus
 ORGANISM Mesocricetus auratus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Cricetidae; Cricetinae; Mesocricetus.
 1
 Nakamura,N., Yoshimi,T. and Miura,T.
 Increased gene expression of lung marker proteins in the homeobox B3-overexpressed fetal lung cell line M3E3/C3
 Cell Growth Differ. 13 (4), 195-203 (2002)
 11971819
 JOURNAL
 PUBLISHED
 11971819
 REFERENCE
 AUTHORS
 TITLE
 Nakamura,N. and Miura,T.
 Direct Submission

JOURNAL Submitted (01-MAR-2001) Nobuatsu Nakamura, Tokyo University of Pharmacy and Life Science, Department of Environmental Molecular Physiology; Horinouchi 1432-1, Hachiohji, Tokyo 192-0392, Japan (E-mail: s94804@educ.its.toyaku.ac.jp, URL: http://www.its.toyaku.ac.jp/Life-Science/lemp-4/, Tel: 81-426-76-5053, Fax: 81-426-76-6811)

FEATURES

source

1..1275
/location/Qualifiers
/organism="Mesocricetus auratus"
/mol_type="mRNA"
/db_xref="taxon:10036"
/cell_line="M3E3/C3"

gene

1..1275
/gene="hox b3"

CDS

1..1275
/gene="hox b3"
/codon_start=1
/product="homeobox B3"
/protein_id="BAB33375.1"
/db_xref="GI:13359293"
/translation="MQKATYYNTAAALFGGYSSYPGNGFCYDGPFPQPPQAATHLE GDYRSACISQISGNAAPHAKEKELNGSCMRPLAPPELPAPGSPPPSAAPTSTIN SNGGSGSKGPKCGGSGNSTLTQIFPMWKESTOTKLKNSPFTGCGGGGGGGG SGGSSGGGGGGGGKSPGSAASKRARTAYTSAQLVELEKEFHFNYLCRRPRVE MANLLLSERQIKIWFQNRMKYKQOKAKGLASSGGSPAGSPQPMQSTAGFMNA LHSMTPSYDSPSPKQGHONAYALPSNVOPPLKGCAGAPQKCPPTPASEVEPHVLOA NGGAYGTPTMQGSFVYGGGVADPLPPAGPSLYGLNHLSHHPSGNDLYNGAAPMGF NQHHGCDPHTTIDLSHHPQGRQIAPN"

misc_feature

385..402
/gene="hox b3"

misc_feature

556..738
/note="hexa peptide"
/gene="hox b3"
/note="homeodomain"

ORIGIN

Query Match 0.7%; Score 23; DB 6; Length 1275;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 GCCGCTGCTACTGCGCGCTGC 48
|||||

Db 501 GCCGCTGCTACTGCGCGCTGC 479
|||||

RESULT 97

AC149896 74579 bp DNA linear VRT 02-SEP-2005
LOCUS Xenopus tropicalis clone ISB-27807, complete sequence.
DEFINITION AC149896
ACCESSION AC149896
VERSION AC149896.2 GI:74096560
KEYWORDS HTG.
SOURCE Xenopus tropicalis (Silurana tropicalis)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL
REFERENCE 2 (bases 1 to 74579)
AUTHORS DOE Joint Genome Institute
TITLE Direct Submission
JOURNAL
REFERENCE 3 (bases 1 to 74579)
AUTHORS Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL
Submitted (02-SEP-2005) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Sep 2, 2005 this sequence version replaced gi:49170153. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center www.shgc.stanford.edu
Quality: Phrap Quality >=40 98.8% of Sequence;
Estimated Total Number of Errors is 0.5.

FEATURES

source

1..74579
/location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB-27807"

ORIGIN

Query Match 0.7%; Score 23; DB 11; Length 74579;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2526 TGATACAGATACACGGAACGTT 2548
|||||

Db 55320 TGATACAGATACACGGAACGTT 55298
|||||

RESULT 98

AC149892 98359 bp DNA linear HTG 24-JUN-2004
LOCUS Xenopus tropicalis clone ISB-242D2, WORKING DRAFT SEQUENCE, 8
DEFINITION AC149892
ACCESSION AC149892
VERSION AC149892.1 GI:49170149
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Xenopus tropicalis (Silurana tropicalis)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE

AUTHORS DOE Joint Genome Institute.
TITLE Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 98359)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL
Submitted (24-JUN-2004) Production Genomics Facility, DOE Joint Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

COMMENT

-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 2865953
Center clone name: ISB-242D2

Summary Statistics
Consensus quality: 95112 bases at least Q40
Consensus quality: 96157 bases at least Q30
Consensus quality: 96888 bases at least Q20
Estimated insert size: 104000; agarose-fp estimation
Estimated insert size: 97659; sum-of-contigs estimation
Quality coverage: 15.53 in Q20 bases; agarose-fp estimation
Quality coverage: 16.54 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.

1 2277: contig of 2277 bp in length
2278: gap of unknown length
2378 5120: contig of 2743 bp in length
5121 5220: gap of unknown length
5221 8464: contig of 3244 bp in length
8465 8564: gap of unknown length
8565 19957: contig of 11392 bp in length
19957 20056: gap of unknown length
20057 30478: contig of 10422 bp in length
30479 30578: gap of unknown length
30579 41701: contig of 11122 bp in length
41701 52926: gap of unknown length
52926 53025: contig of 11125 bp in length
53026 98359: gap of unknown length
53026 98359: contig of 45334 bp in length.

FEATURES

source
1..98359
location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB-242D2"
2278..2377
/estimated_length=unknown
5121..5220
/estimated_length=unknown
8465..8564
/estimated_length=unknown
19957..20056
/estimated_length=unknown
30479..30578
/estimated_length=unknown
41701..41800
/estimated_length=unknown
52926..53025
/estimated_length=unknown

ORIGIN

Query Match 0.7%; Score 23; DB 12; Length 98359;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2526 TGATACAGGATACACGGACGTT 2548
Db 55017 TGATACAGGATACACGGACGTT 55039

RESULT 99
AC161463 213261 bp DNA linear VRT 16-DEC-2005
LOCUS Gallus gallus BAC clone CH261-127J9 from chromosome ul, complete
DEFINITION sequence.

AC161463
AC161463.3 GI:83699520
ACCESSION HTG.
KEYWORDS Gallus gallus (chicken)
SOURCE
ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus

1 (bases 1 to 213261)
Delgado B, Kozlowski A, and Lewis S.
The sequence of Gallus gallus BAC clone CH261-127J9
Unpublished (2001)

2 (bases 1 to 213261)
Wilson R.K.

AUTHORS
TITLE Direct Submission

JOURNAL Submitted (14-MAY-2005) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 213261)

REFERENCE
AUTHORS Wilson, R.K.
TITLE Direct Submission

JOURNAL

REFERENCE Submitted (29-OCT-2005) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 213261)

AUTHORS Wilson, R.K.

TITLE Direct Submission

JOURNAL Submitted (16-DEC-2005) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA

COMMENT On Dec 16 2005 this sequence version replaced gi:78272069.

-----Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: submissions@wustl.edu

-----Summary Statistics

Center project name: J_AA127J09

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

SOURCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES
source
1..213261
location/Qualifiers
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
/chromosome="ul"
/clone="CH261-127J9"
/clone_lib="CHORI261"

ORIGIN

Query Match 0.7%; Score 23; DB 11; Length 213261;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1477 ATCTATGAAGAAACAACAGACAT 1499
Db 25695 ATCTATGAAGAAACAACAGACAT 25717

RESULT 100

AR631303

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

-----Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: submissions@wustl.edu

-----Summary Statistics

Center project name: J_AA127J09

Query Match 0.7%; Score 22; DB 2; Length 22;

Thu Jun 22 09:04:25 2006

```

RESULT 101
LOCUS AX608772/c
DEFINITION Sequence 49 from Patent WO0231134.
ACCESSION AX608772
VERSION AX608772.1 GI:28404324
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
TITLE Novel serine protease genes related to DPPIV
JOURNAL Patent: WO 0231134-A 49 18-APR-2002;
Ferring BV (NL)
FEATURES
SOURCE Location/Qualifiers
1. .22
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 0.7%; Score 22; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.7; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;
QY 1972 TCAAGTCCTGAAGATGACCCAA 1993
DB 22 TCAAGTCCTGAAGATGACCCAA 1
RESULT 104
LOCUS AR631301/c
DEFINITION Sequence 46 from patent US 6844180.
ACCESSION AR631301
VERSION AR631301.1 GI:59770960
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
AUTHORS Qi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 46 18-JAN-2005;
Ferring BV;;
FEATURES
SOURCE Location/Qualifiers
1. .39
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 22; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 4.8; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;
QY 2839 ATTGCTGCTCTAAAGTGCATAT 2860
DB 39 ATTGCTGCTCTAAAGTGCATAT 18
RESULT 105
LOCUS AX608769/c
DEFINITION Sequence 46 from Patent WO0231134.
ACCESSION AX608769
VERSION AX608769.1 GI:28404321
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

```

```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
TITLE Novel serine protease genes related to dppiv
JOURNAL Patent: WO 0231134-A 46 18-APR-2002,
Ferring BV (NL)
FEATURES
source Location/Qualifiers
1..39
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 0.7%; Score 22; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2839 ATTGCTGCTCTAAAGTGATAT 2860
Db 39 ATTGCTGCTCTAAAGTGATAT 18
RESULT 106
CT030588 2250 bp mRNA linear VRT 12-OCT-2005
LOCUS CT030588
DEFINITION Xenopus tropicalis finished cDNA, clone TTPA012113.
ACCESSION CT030588
VERSION CT030588.1 GI:77623488
KEYWORDS Xenopus tropicalis (Silurana tropicalis)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 2250)
REFERENCE
AUTHORS Anaya, E., Ashurst, J.L., Bonfield, J.K., Croning, M.D.R., Davies, R.M.,
Francis, M.D., Garrett, N., Gilchrist, M.J., Graham, D.V.,
McLaren, S.R., Papalopulu, N., Rogers, J., Smith, J.C., Taylor, R.G.,
Voigt, J. and Zorn, A.M.
TITLE Direct Submission
JOURNAL Submitted (12-OCT-2005) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
COMMENT Sanger Xenopus tropicalis EST/cDNA project.
This sequence is from a Xenopus Gene Collection (XGC) library, from
a library constructed by Nigel Garrett. cDNA was prepared from RNA
extracted from tadpoles. EcoRI-NotI cut cDNA was then ligated into
pCS107 with EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..2250
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TTPA012113"
/dev_stage="tadpole (stage 35-40)"
ORIGIN
Query Match 0.7%; Score 22; DB 11; Length 2250;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1073 ATGAGAAATGATGATCTGA 1094
Db 167 ATGAGAAATGATGATCTGA 188
RESULT 107
BD103383 3322 bp DNA linear PAT 27-AUG-2002
LOCUS BD103383
DEFINITION Novel TTV-associated virus s-TTV.

```

```

BD103383
KEYWORDS WO 0185771-A/9.
SOURCE unidentified
ORGANISM unidentified
REFERENCE
1 (bases 1 to 3322)
unclassified sequences.
AUTHORS Abe, K.
TITLE Novel TTV-associated virus s-TTV
JOURNAL Patent: WO 0185771-A 9 15-NOV-2001;
EISAI CO LTD, KENJI ABE
COMMENT OS TTV CH71
PN WO 0185771-A/9
PD 15-NOV-2001
PF 11-MAY-2001 WO 2001JP003954
PR 11-MAY-2000 JP OOP 137894, 08-NOV-2000 JP OOP 340614 PI
KENJI ABE
PC C07K14/01, C12N15/51, G01N33/576, G01N33/569, C12Q1/68//C12P21/02
CC Novel TTV-associated virus s-TTV
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..3322
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 0.7%; Score 22; DB 2; Length 3322;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28 CGCTGCTACTGCCGCCGCTGCT 49
Db 115 CGCTGCTACTGCCGCCGCTGCT 136
RESULT 108
BD103384 3322 bp DNA linear PAT 27-AUG-2002
LOCUS BD103384
DEFINITION Novel TTV-associated virus s-TTV.
ACCESSION BD103384
VERSION BD103384.1 GI:22648958
KEYWORDS WO 0185771-A/10.
SOURCE unidentified
ORGANISM unidentified
REFERENCE
1 (bases 1 to 3322)
unclassified sequences.
AUTHORS Abe, K.
TITLE Novel TTV-associated virus s-TTV
JOURNAL Patent: WO 0185771-A 10 15-NOV-2001;
EISAI CO LTD, KENJI ABE
COMMENT OS TTV CH71
PN WO 0185771-A/10
PD 15-NOV-2001
PF 11-MAY-2001 WO 2001JP003954
PR 11-MAY-2000 JP OOP 137894, 08-NOV-2000 JP OOP 340614 PI
KENJI ABE
PC C07K14/01, C12N15/51, G01N33/576, G01N33/569, C12Q1/68//C12P21/02
CC Novel TTV-associated virus s-TTV
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..3322
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 0.7%; Score 22; DB 2; Length 3322;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Qy	28	CGCTGCTACTGCGCGCGTGT	49
Db	115	CGCTGCTACTGCGCGCGTGT	136
RESULT 109			
AB049608	3322 bp	DNA	linear VRL 10-JAN-2001
LOCUS	TT virus gene for ORF1, ORF2, complete cds, isolate:s-TTV CH71.		
DEFINITION	AB049608		
ACCESSION	AB049608.1		
VERSION	GI:12060430		
KEYWORDS	Torque teno virus (TTV)		
SOURCE	Torque teno virus		
ORGANISM	Viruses; ssDNA viruses; Anellovirus.		
REFERENCE	1 (sites)		
AUTHORS	Abe,K. and Inami,T.		
TITLE	simian TTV		
JOURNAL	Unpublished		
REFERENCE	2 (sites)		
AUTHORS	Abe,K. and Inami,T.		
TITLE	simian TT virus		
JOURNAL	Published Only in DataBase (2001)		
REFERENCE	3 (bases 1 to 3322)		
AUTHORS	Abe,K. and Inami,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-OCT-2000) Kenji Abe, National Institute of Infectious Diseases, Department of Pathology; Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan (E-mail:kenjiabe@nih.go.jp, Tel:81-3-5285-1111(ex.2624), Fax:81-3-5285-1189)		
Location/Qualifiers			
FEATURES	1..3322		
source	/organism="Torque teno virus"		
	/mol_type="genomic DNA"		
	/isolate="s-TTV CH71"		
	/specific_host="pan troglodytes"		
	/db_xref="taxon:68887"		
	153..518		
CDS	/codon_start=1		
	/product="ORF2"		
	/protein_id="BAB20603.1"		
	/db_xref="GI:12060431"		
	/translation="MSLWRPPAHVDVTGLERNWYSCFRSHAMCGGNFIDLHLNLAS		
	RFGRRPPRPPGAPQPPPIPLPALPAPEGPRAPWRSGGGGDRGGDAPRRGGD		
	DGDGLADAEALLAELEDDAE"		
	389..2572		
CDS	/codon_start=1		
	/product="ORF1"		
	/protein_id="BAB20604.1"		
	/db_xref="GI:12060432"		
	/translation="MAFWRRRRRRARRRRRAAARRRRRRRRRRRAARRRRRRRVR		
	RRRRRRARRRRRRRRRRRRRRKRLIIRQWQPYTRRCYIVGLPLIMCGENCFSKN		
	FATHSDIDCSVAGGTMATQFTLRLYDEFOHLNYWTASNNDLDLARYLGATIT		
	FRPDPDVFVQIHTSPPRTEMTGSIHPGMLKKHRLIPSLKTRPGRKHRTVR		
	IGAPKTFEDVQPTDCLDVLVYATADTFPPGSLPLNTHCVNFQVLGSDYNDI		
	LSILPDKLKSKEIKIYKINAIYNTKQITIAHLKIDNCKNTQTDINTKHNTOYK		
	GNTVVDINQVRDGTGDKFKATSLASNSWVKPTNDLEVHTGMYSSIFLSSGRNP		
	ELKGPYTDVCPNLDVGIGNVLWDCSKDSKFDTSKVPYIRDMPLWAALFGYAE		
	YASKEGTDTAINTNARTLICPYTDMPLFKPQDPKFGFVPSYLNFGLMFGDSTYVP		
	IRQAKWYACLYHQQVIEAITQSGPPFVHCDYKSAVLGVKYKFKWIGGSPFIHQVV		
	RNPCTHSTQGRKPRSVQVTRKYNTPETFTWSDTRGLFGDGLQRLMOQPL		
	NELPYTAKRPRKRDTLAEHQHEPEPDVSDFKRLRLQLOQQPWLDSQBEKKARESKA		
	QEEAKTLEOQLEQQRLLGFQLRCLAYQVVLQVQKQGHAIIPQALVQ"		
ORIGIN			
Query Match	0.7%;	Score 22;	DB 10; Length 3322;
Best Local Similarity	100.0%;	Pred. No. 5.3;	
Matches	22;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	28	CGCTGCTACTGCGCGCGTGT	49
Db	115	CGCTGCTACTGCGCGCGTGT	136
RESULT 110			
HSDJ87H18	43049 bp	DNA	linear PRI 18-MAY-2005
LOCUS	Human DNA sequence from clone RPI-87H18 on chromosome 6q22.1-22.32		
DEFINITION	Contains STSs and GSSs, complete sequence.		
ACCESSION	AL121846		
VERSION	AL121846.6		
KEYWORDS	GI:6630794		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumetazoa; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 43049)		
TITLE	Loveall,J.		
JOURNAL	Direct Submission		
COMMENT	Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Dec 22, 1999 this sequence version replaced gi:6599100. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6 RPI-87H18 is from the library RPCI-1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pCYPAC2 ----- Genome Center Center: Wellcome Trust Sanger Institute Center code: SC Web site: http://www.sanger.ac.uk Contact: vegas@sanger.ac.uk ----- This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.		
Location/Qualifiers			
FEATURES	1..43049		
source	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="RZPD:RCP1P704H1887"		
	/db_xref="taxon:9606"		
	/chromosome="6"		
	/map="q22.1-22.32"		
	/clone="RPI-87H18"		
	/clone_lib="RPCI-1"		
	misc_feature		
	1		
	misc_feature		
	43049		
	/note="Clone_left_end: RPI-87H18"		
	/note="Clone_right_end: RPI-87H18"		
ORIGIN			
Query Match	0.7%;	Score 22;	DB 5; Length 43049;
Best Local Similarity	100.0%;	Pred. No. 5.6;	
Matches	22;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	3099	TAAGATTACTAAAAAAA	3120
Db	3402	TAAGATTACTAAAAAAA	3423
RESULT 111			

```

AP000861/c
LOCUS
DEFINITION Homo sapiens chromosome 11 clone CMB9-49N2 map liq25, WORKING DRAFT
SEQUENCE, 14 unordered pieces.
ACCESSION AP000861
VERSION AP000861.3 GI:8119009
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo (104012)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.,
Homo sapiens 104,012 genomic DNA of liq25
Published Only in DataBase (1999)
2 (Bases 1 to 104012)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.,
Direct Submission
Submitted (13-DEC-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
On May 31, 2000 this sequence version replaced gi:7242895.
----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: CMB9-49N2
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 91543 bases at least Q40
Consensus quality: 97504 bases at least Q50
Consensus quality: 100932 bases at least Q20
Insert size: 102712; sum-of-contigs
Quality coverage: 4.48x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
14 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.
1 20152 contig of 20152 bp in length
20253 33930 contig of 13678 bp in length
46313 58306 contig of 12182 bp in length
58407 68225 contig of 11994 bp in length
68326 74172 contig of 5847 bp in length
74273 80325 contig of 6053 bp in length
80426 85720 contig of 5295 bp in length
85821 89487 contig of 3667 bp in length
89588 93995 contig of 4408 bp in length
94096 97810 contig of 3715 bp in length
97911 100367 contig of 4408 bp in length
100468 102265 contig of 2457 bp in length
102366 104012 contig of 1798 bp in length
Sequence updated (13-Mar-2000)
Sequence updated (26-May-2000).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 20152: contig of 20152 bp in length
20253: gap of 100 bp
33930: contig of 13678 bp in length
34031: gap of 100 bp
46313: contig of 12182 bp in length
46313: gap of 100 bp
58306: contig of 11994 bp in length
58407: gap of 100 bp
68225: contig of 5847 bp in length
68326: gap of 100 bp
74172: contig of 5847 bp in length
74273: gap of 100 bp
80325: contig of 6053 bp in length
80426: gap of 100 bp
85720: contig of 5295 bp in length
85821: gap of 100 bp
89487: contig of 3667 bp in length
89588: gap of 100 bp
93995: contig of 4408 bp in length
94096: gap of 100 bp
97810: contig of 3715 bp in length
97911: gap of 100 bp
100367: contig of 2457 bp in length
100468: gap of 100 bp
102265: contig of 1798 bp in length
102366: gap of 100 bp
102366: 104012: contig of 1647 bp in length.
FEATURES
Location/Qualifiers
source
1. 104012
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11Q25"
/clone="CMB9-49N2"
1. 20152
/note="assembly_fragment"
misc_feature
20253..33930
/note="assembly_fragment"
misc_feature
34031..46212
/note="assembly_fragment"
misc_feature
46313..58306
/note="assembly_fragment"
misc_feature
58407..68225
/note="assembly_fragment"
misc_feature
68326..74172
/note="assembly_fragment clone_end:T7 vector_side:right"
misc_feature
74273..80325
/note="assembly_fragment"
misc_feature
80426..85720
/note="assembly_fragment"
misc_feature
85821..89487
/note="assembly_fragment"
misc_feature
89588..93995
/note="assembly_fragment"
misc_feature
94096..97810
/note="assembly_fragment"
misc_feature
97911..100367
/note="assembly_fragment"
misc_feature
100468..102265
/note="assembly_fragment"
misc_feature
102366..104012
/note="assembly_fragment"
ORIGIN
Query Match 0.7%; Score 22; DB 12; Length 104012;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

* consists of 11 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 17900: contig of 17900 bp in length
 18000: gap of 100 bp
 18001 34310: contig of 16310 bp in length
 34311 34410: gap of 100 bp
 34411 48225: contig of 13815 bp in length
 48226 48325: gap of 100 bp
 48326 62336: contig of 14011 bp in length
 62337 62436: gap of 100 bp
 62437 76488: contig of 14052 bp in length
 76489 76588: gap of 100 bp
 76589 84448: contig of 7860 bp in length
 84449 84548: gap of 100 bp
 84549 92441: contig of 7893 bp in length
 92442 92541: gap of 100 bp
 92542 99422: contig of 6881 bp in length
 99423 103359: contig of 3837 bp in length
 103360 103459: gap of 100 bp
 103460 104651: contig of 1192 bp in length
 104652 104751: gap of 100 bp
 104752 105350: contig of 599 bp in length.

FEATURES

Location/Qualifiers
 1. .105350
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q25"
 /clone="CMB9-2G15"
 1. .17900
 /note="assembly_fragment"
 18001. .34310
 /note="assembly_fragment"
 34411. .48225
 /note="assembly_fragment"
 48326. .62336
 /note="assembly_fragment"
 62437. .76488
 /note="assembly_fragment"
 76589. .84448
 /note="assembly_fragment"
 84549. .92441
 /note="assembly_fragment"
 92542. .99422
 /note="assembly_fragment clone_end:T7 vector_side:right"
 99523. .103359
 /note="assembly_fragment"
 103460. .104651
 /note="assembly_fragment"
 104752. .105350
 /note="assembly_fragment clone_end:SP6 vector_side:left"

ORIGIN

Query Match 0.7%; Score 22; DB 12; Length 105350;
 Best Local Similarity 100.0%; Pred.No. 5.7;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3098 ATAAGAAATTACTAAAAAAA 3119
 Db 88539 ATAAGAAATTACTAAAAAAA 88518

RESULT 113
 AP008208_040/c
 WPCOMMENT

Sequence split into 360 fragments LOCUS AP008208 Accession AP008208

QY 3098 ATAAGAAATTACTAAAAAAA 3119
 Db 84413 ATAAGAAATTACTAAAAAAA 84392

RESULT 112
 AP000661/c 105350 bp DNA linear HTG 30-MAY-2000
 LOCUS Homo sapiens chromosome 11 clone CMB9-2G15 map 11q25, WORKING DRAFT
 DEFINITION SEQUENCE, 11 unordered pieces.

ACCESSION AP000661
 VERSION HTG; HTGS PHASE1; HTGS_DRAFT.

KEYWORDS Homo sapiens (human)

SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE
 AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE Homo sapiens 105,350 genomic DNA of 11q25

JOURNAL Published Only in Database (1999)

REFERENCE
 AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE Direct Submission

JOURNAL Submitted (04-NOV-1999) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
 Japan [E-mail:hattori@gsc.riken.go.jp,
 URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
 Fax:81-42-778-9924]
 On May 31, 2000 this sequence version replaced gi:6997535.

COMMENT

----- Genome Center
 Center: RIKEN Genomic Sciences Center (GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsc.riken.go.jp
 ----- Project Information
 Center project name: HumDraft11
 Center clone name: CMB9-2G15
 ----- Summary Statistics
 Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 101290 bases at least Q40
 Consensus quality: 103227 bases at least Q30
 Consensus quality: 103986 bases at least Q20
 Insert size: 104350; sum-of-contigs
 Quality coverage: 8.12x in Q20 bases; sum-of-contigs

 NOTE: This is a 'working draft' sequence. It currently consists of
 11 contigs. The true order of the pieces is not known and their
 order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved

1 17900 contig of 17900 bp in length
 18001 34310 contig of 16310 bp in length
 34411 48225 contig of 13815 bp in length
 48326 62336 contig of 14011 bp in length
 62437 76488 contig of 14052 bp in length
 76589 84448 contig of 7860 bp in length
 84549 92441 contig of 7893 bp in length
 92542 99422 contig of 6881 bp in length
 99523 103359 contig of 3837 bp in length
 103460 104651 contig of 1192 bp in length
 104752 105350 contig of 599 bp in length

Sequence updated (06-Jan-2000)

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently

Fragment Name	Begin	End			
AP008208_000	1	110000	AP008208_072	7200001	7310000
AP008208_001	100001	210000	AP008208_073	7300001	7410000
AP008208_002	200001	310000	AP008208_074	7400001	7510000
AP008208_003	300001	410000	AP008208_075	7500001	7610000
AP008208_004	400001	510000	AP008208_076	7600001	7710000
AP008208_005	500001	610000	AP008208_077	7700001	7810000
AP008208_006	600001	710000	AP008208_078	7800001	7910000
AP008208_007	700001	810000	AP008208_079	7900001	8010000
AP008208_008	800001	910000	AP008208_080	8000001	8110000
AP008208_009	900001	1010000	AP008208_081	8100001	8210000
AP008208_010	1000001	1110000	AP008208_082	8200001	8310000
AP008208_011	1100001	1210000	AP008208_083	8300001	8410000
AP008208_012	1200001	1310000	AP008208_084	8400001	8510000
AP008208_013	1300001	1410000	AP008208_085	8500001	8610000
AP008208_014	1400001	1510000	AP008208_086	8600001	8710000
AP008208_015	1500001	1610000	AP008208_087	8700001	8810000
AP008208_016	1600001	1710000	AP008208_088	8800001	8910000
AP008208_017	1700001	1810000	AP008208_089	8900001	9010000
AP008208_018	1800001	1910000	AP008208_090	9000001	9110000
AP008208_019	1900001	2010000	AP008208_091	9100001	9210000
AP008208_020	2000001	2110000	AP008208_092	9200001	9310000
AP008208_021	2100001	2210000	AP008208_093	9300001	9410000
AP008208_022	2200001	2310000	AP008208_094	9400001	9510000
AP008208_023	2300001	2410000	AP008208_095	9500001	9610000
AP008208_024	2400001	2510000	AP008208_096	9600001	9710000
AP008208_025	2500001	2610000	AP008208_097	9700001	9810000
AP008208_026	2600001	2710000	AP008208_098	9800001	9910000
AP008208_027	2700001	2810000	AP008208_099	9900001	10010000
AP008208_028	2800001	2910000	AP008208_100	10000001	10110000
AP008208_029	2900001	3010000	AP008208_101	10100001	10210000
AP008208_030	3000001	3110000	AP008208_102	10200001	10310000
AP008208_031	3100001	3210000	AP008208_103	10300001	10410000
AP008208_032	3200001	3310000	AP008208_104	10400001	10510000
AP008208_033	3300001	3410000	AP008208_105	10500001	10610000
AP008208_034	3400001	3510000	AP008208_106	10600001	10710000
AP008208_035	3500001	3610000	AP008208_107	10700001	10810000
AP008208_036	3600001	3710000	AP008208_108	10800001	10910000
AP008208_037	3700001	3810000	AP008208_109	10900001	11010000
AP008208_038	3800001	3910000	AP008208_110	11000001	1110000
AP008208_039	3900001	4010000	AP008208_111	11100001	11210000
AP008208_040	4000001	4110000	AP008208_112	11200001	11310000
AP008208_041	4100001	4210000	AP008208_113	11300001	11410000
AP008208_042	4200001	4310000	AP008208_114	11400001	11510000
AP008208_043	4300001	4410000	AP008208_115	11500001	11610000
AP008208_044	4400001	4510000	AP008208_116	11600001	11710000
AP008208_045	4500001	4610000	AP008208_117	11700001	11810000
AP008208_046	4600001	4710000	AP008208_118	11800001	11910000
AP008208_047	4700001	4810000	AP008208_119	11900001	12010000
AP008208_048	4800001	4910000	AP008208_120	12000001	12110000
AP008208_049	4900001	5010000	AP008208_121	12100001	12210000
AP008208_050	5000001	5110000	AP008208_122	12200001	12310000
AP008208_051	5100001	5210000	AP008208_123	12300001	12410000
AP008208_052	5200001	5310000	AP008208_124	12400001	12510000
AP008208_053	5300001	5410000	AP008208_125	12500001	12610000
AP008208_054	5400001	5510000	AP008208_126	12600001	12710000
AP008208_055	5500001	5610000	AP008208_127	12700001	12810000
AP008208_056	5600001	5710000	AP008208_128	12800001	12910000
AP008208_057	5700001	5810000	AP008208_129	12900001	13010000
AP008208_058	5800001	5910000	AP008208_130	13000001	13110000
AP008208_059	5900001	6010000	AP008208_131	13100001	13210000
AP008208_060	6000001	6110000	AP008208_132	13200001	13310000
AP008208_061	6100001	6210000	AP008208_133	13300001	13410000
AP008208_062	6200001	6310000	AP008208_134	13400001	13510000
AP008208_063	6300001	6410000	AP008208_135	13500001	13610000
AP008208_064	6400001	6510000	AP008208_136	13600001	13710000
AP008208_065	6500001	6610000	AP008208_137	13700001	13810000
AP008208_066	6600001	6710000	AP008208_138	13800001	13910000
AP008208_067	6700001	6810000	AP008208_139	13900001	14010000
AP008208_068	6800001	6910000	AP008208_140	14000001	14110000
AP008208_069	6900001	7010000	AP008208_141	14100001	14210000
AP008208_070	7000001	7110000	AP008208_142	14200001	14310000
AP008208_071	7100001	7210000	AP008208_143	14300001	14410000
			AP008208_144	14400001	14510000

```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
Published Only in Database (2000)
TITLE
JOURNAL
AUTHORS
2 (Bases 1 to 117345)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (07-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@psc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Nov 5, 2001 this sequence version replaced gi:11138088.
COMMENT
FEATURES
source
1..117345
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-700F16"

ORIGIN
Query Match 0.7%; Score 22; DB 5; Length 117345;
Best Local Similarity 100.0%; Pred. No. 5.7; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

Qy 3098 ATAGAATTACTTAAAAA 3119
Db 4058 ATAGAATTACTTAAAAA 4079

RESULT 115
AL160263 117445 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome 6 clone RP1-206M12, 22 unordered pieces.
DEFINITION
ACCESSION AL160263
VERSION AL160263.6 GI:9863592
KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS
Sims,S.
TITLE
JOURNAL
Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:8977651.
COMMENT
1
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: du206M12
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: M13; M7815; 4% of reads
Sequencing vector: plasmid; 108752; 95% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 104927 bases at least Q40
Consensus quality: 109617 bases at least Q30
Consensus quality: 112114 bases at least Q20
Insert size: 115345; sum-of-contigs
Insert size: 179782; 18.8% error; agarose-fp
Quality coverage: 4.71x in Q20 bases; sum-of-contigs Quality

14500001 14610000
14600001 14710000
14700001 14810000
14800001 14910000
14900001 15010000
15000001 15110000
15100001 15210000
15200001 15310000
15300001 15410000
15400001 15510000
15500001 15610000
15600001 15710000
15700001 15810000
15800001 15910000
15900001 16010000
16000001 16110000
16100001 16210000
16200001 16310000
16300001 16410000
16400001 16510000
16500001 16610000
16600001 16710000
16700001 16810000
16800001 16910000
16900001 17010000
17000001 17110000
17100001 17210000
17200001 17310000
17300001 17410000
17400001 17510000
17500001 17610000
17600001 17710000
17700001 17810000
17800001 17910000
17900001 18010000
18000001 18110000
18100001 18210000
18200001 18310000
18300001 18410000
18400001 18510000
18500001 18610000
18600001 18710000
18700001 18810000
18800001 18910000
18900001 19010000
19000001 19110000
19100001 19210000
19200001 19310000
19300001 19410000
19400001 19510000
19500001 19610000
19600001 19710000

Query Match 0.7%; Score 22; DB 4; Length 110000;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 986 AAGAAGATTGTGATGATATTC 1007
Db 46307 AAGAAGATTGTGATGATATTC 46286

RESULT 114
AP001775
LOCUS Homo sapiens genomic DNA, chromosome 11q clone:RP11-700F16,
DEFINITION complete sequence.
ACCESSION AP001775
VERSION AP001775.4 GI:16751482
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

coverage: 3.68x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence, as soon as it is available and the accession number will be preserved.

1 4694: contig of 4694 bp in length
4695 4794: gap of 100 bp
4795 7828: contig of 3034 bp in length
7829 7928: gap of 100 bp
7929 10684: contig of 2756 bp in length
10685 10784: gap of 100 bp
10785 13948: contig of 3164 bp in length
13949 14048: gap of 100 bp
14049 17288: contig of 3240 bp in length
17289 17388: gap of 100 bp
17389 21695: contig of 4307 bp in length
21696 21795: gap of 100 bp
21796 24617: contig of 2822 bp in length
24618 29336: contig of 4619 bp in length
29337 33478: contig of 4042 bp in length
33479 33578: gap of 100 bp
33579 36895: contig of 3317 bp in length
36896 36995: gap of 100 bp
36996 39561: contig of 2566 bp in length
39562 44841: gap of 100 bp
44842 47261: contig of 5080 bp in length
47262 47361: contig of 2420 bp in length
47362 50478: gap of 100 bp
50479 50578: contig of 3317 bp in length
50579 93459: gap of 100 bp
93460 93559: gap of 100 bp
93560 96033: contig of 2474 bp in length
96034 96133: gap of 100 bp
96134 100979: contig of 4846 bp in length
100980 101079: gap of 100 bp
101080 103587: contig of 2508 bp in length
103588 106442: contig of 2755 bp in length
106443 106543: gap of 100 bp
106544 109973: contig of 3431 bp in length
109974 110074: gap of 100 bp
110075 114633: contig of 4561 bp in length
114634 114735: gap of 100 bp
114735 117445: contig of 2711 bp in length.

FEATURES

source
1..117445
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP1-206M12"
/clone_lib="RPC1-1"
1..4694
note="assembly_fragment:00720
fragment_chain:1"
4795..7828
note="assembly_fragment:01211
fragment_chain:1"
7929..10684
note="assembly_fragment:00482
fragment_chain:1"
10785..13948
note="assembly_fragment:01140
fragment_chain:1"

misc_feature 14049..17288
/note="assembly_fragment:00231
fragment_chain:2"
misc_feature 17389..21695
/note="assembly_fragment:00129
fragment_chain:2"
misc_feature 21796..24617
/note="assembly_fragment:01301
fragment_chain:3"
misc_feature 24718..29336
/note="assembly_fragment:00659
fragment_chain:3"
misc_feature 29437..33478
/note="assembly_fragment:01946
fragment_chain:4"
misc_feature 33579..36895
/note="assembly_fragment:00111
fragment_chain:4"
misc_feature 36996..39561
/note="assembly_fragment:00205"
fragment_chain:4"
misc_feature 39662..44741
/note="assembly_fragment:00242"
fragment_chain:4"
misc_feature 44842..47261
/note="assembly_fragment:00673"
fragment_chain:4"
misc_feature 47362..50478
/note="assembly_fragment:00870"
fragment_chain:4"
misc_feature 50579..50578
/note="assembly_fragment:01072"
fragment_chain:4"
misc_feature 50579..93459
/note="assembly_fragment:01200"
fragment_chain:4"
misc_feature 93560..96033
/note="assembly_fragment:01517"
fragment_chain:4"
misc_feature 96134..100979
/note="assembly_fragment:01557"
fragment_chain:4"
misc_feature 100980..103587
/note="assembly_fragment:01608"
fragment_chain:4"
misc_feature 103588..106442
/note="assembly_fragment:01719"
fragment_chain:4"
misc_feature 106543..109973
/note="assembly_fragment:01865"
fragment_chain:4"
misc_feature 110074..114633
/note="assembly_fragment:01866"
fragment_chain:4"

ORIGIN

Query Match 0.7%; Score 22; DB 12; Length 117445;
Best Local Similarity 100.0%; Pred. No. 5.7; 0; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0;

Qy 3099 TAGAATTACTAAAAA 3120
|||||
Db 52142 TAGAATTACTAAAAA 52163
|||||

RESULT 116

CR749173
LOCUS CR749173 143028 bp DNA linear HTG 03-MAR-2005
DEFINITION Danio rerio clone DKEYP-73C1, 5 unordered pieces.
ACCESSION CR749173
VERSION CR749173.10 GI:60495570
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 Phil James et al. 143028
AUTHORS Phil James, B. J. Harrison
TITLE Submitted (03-MAR-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
JOURNAL On Mar 3 2005 this sequence version replaced gi:56783567.
COMMENT ----- Genome Center

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfsh-help@sanger.ac.uk
----- Project Information
Center project name: zkp73C1
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 140742 bases at least Q40
Consensus quality: 141127 bases at least Q30
Consensus quality: 141713 bases at least Q20
Insert size: 142628; sum-of-contigs
Insert size: 196437; 2.5% error; agarose-fp
Quality coverage: 7.92x in Q20 bases; sum-of-contigs Quality
Coverage: 6.95x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 123412: contig of 123412 bp in length
* 123413 123512: gap of 100 bp
* 123513 135554: contig of 12042 bp in length
* 135554 135655: gap of 100 bp
* 135655 138272: contig of 2618 bp in length
* 138272 138373: gap of 100 bp
* 138373 140636: contig of 2264 bp in length
* 140636 140737: gap of 100 bp
* 140737 143028: contig of 2292 bp in length.
*
* Location/Qualifiers
* 1.143028
* /organism="Danio rerio"
* /mol_type="genomic DNA"
* /db_xref="taxon:7955"
* /clone="DKEYP-73C1"
* /clone_lib="DanioKeyPilot"
*
misc_feature 1.123412
/note="assembly_fragment:00549
fragment_chain:1"
123513..135554
/note="assembly_fragment:01662
fragment_chain:1"
135655..138272
/note="assembly_fragment:01725
fragment_chain:2"
138373..140636
/note="assembly_fragment:01743
fragment_chain:2"
140737..143028
/note="assembly_fragment:01700"

ORIGIN
Query Match 0.7%; Score 22; DB 12; Length 143028;
Best Local Similarity 100.0%; Pred. No. 5.7; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;
QY 3051 ATGCTCAAAATCAATGATACA 3072
|||||
DB 36416 ATGCTCAAAATCAATGATACA 36437
|||||
RESULT 117
CR847930 167262 bp DNA linear VRT 29-JAN-2005
LOCUS Zebrafish DNA sequence from clone DKEY-148H19 in linkage group 19,
DEFINITION complete sequence.
ACCESSION CR847930
VERSION CR847930.5 GI:58331654

HTG
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 167262)
Clark, S.
Direct Submission
Submitted (29-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequests@sanger.ac.uk
zfsh-help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Jan 29, 2005 this sequence version replaced gi:56310288.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfsh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhihong Bao and Sean Eddy, submitted), and those
beginning 'drr' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-148H19
is from a Zebrafish BAC library
VECTOR: pindigobAC-5.

FEATURES
source
1.167262
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-148H19"
/clone_lib="DanioKey"

ORIGIN
Query Match 0.7%; Score 22; DB 11; Length 167262;
Best Local Similarity 100.0%; Pred. No. 5.8; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;
QY 3051 ATGCTCAAAATCAATGATACA 3072
|||||
DB 166541 ATGCTCAAAATCAATGATACA 166562
|||||
RESULT 118
AC158877/c 169540 bp DNA linear HTG 24-MAR-2005
LOCUS Bos taurus clone CH240-22115, WORKING DRAFT SEQUENCE, 23 unordered
DEFINITION pieces.
ACCESSION AC158877

```

VERSION
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
ORGANISM
Bos taurus (cattle)

REFERENCE
AUTHORS
AC158877.1 GI:61740749
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 169540)
Murny D., Adams C., Agbai II O., Allen C., Alsbrooks S., Archer P.,
Arredondo H., Bandaranaike D., Bangura L., Beltran B., Beltran R.,
Beraducci A., Biswal K., Blyth P., Bonham H., Buhay C., Burch P.,
Cadorese I., Candau A., Cardenas V., Carter K., Cavazos I.,
Chacko J., Chaurou M., Chavez D., Chen A., Chen G., Chen R.,
Cheng K.-I., Chu J., Clerc K., Cockrell R., Coyle M., Cree A.,
Curry S., Dai W., Davila M.L., Davis C., Davy-Carroll L., De
Anda C., Delgado O., Denson S., Deramo C., Ding Y., Dinh H.,
Donlin J., McCauley S., Dugan-Rocha S., Dunn A., Durbin K.,
Dziuda D., Egan A., Escotto M., Espinosa V., Eugene C., Fa M.,
Fernandez S., Fernando P., Flagg N., Forbes L., Foster P.,
Fowler G., Fu Q., Fuh E., Garcia A., Garcia R., Garner T.,
Gaskin C., Gench S., Ghose S., Gill R., Gonzalez D.,
Gonzalez-Garay M., Guevara W., Holder M., Haaaland W., Haeblerlen K.,
Hall B., Hamid H., Hamilton K., Harbes B., Harris R., Havlak P.,
Hines S., Hitchens M., Hodgson A., Hogues M., Hollins B.,
Howell L.T., Hulyk S., Hume J., Imo K., Jackson A., Jackson L.,
Jacob L., Jiang H., Johnson B., Johnson R., Kalafus K., Kelly S.,
Keys T., Khan Z., King L., Kovar C., Kowis A., Kowis C., Lara F.,
Leal S., Lee K., Lee S., LeGall F.I., Lemon S., Lewis L., Li B.,
Li Y., Li Z., Linnell M., Liu W., Liu Y.-S., Liu Y., Liyanage D.,
London P., Lopez J., Lorenzuehwa L., Lozano R., Luk T., Madur R.,
Maheshwari M., Mahoney C., Malloy K., Mansouri D., Martinez E.,
McClelland H., McPherson J., Mercadao C., Metzker M.,
Milosavljevic A., Minja E., Morgan M., Morris S.,
Murray D., Nazareth L., Ngo D., Nguyen N., Norwig-Eastaugh E.,
Nott A., Nwaokemele O., Obregon M., Ochi-Okojie C., Oden E.,
Okwuonu G., Okwuonu K., Parker D., Pasternak S., Patel B.,
Patel V., Paul H., Perez A., Perez R., Quim A., Quim F.,
Primus E., Pu L.-L., Puazo M., Qin X., Quim A., Quimoz J.,
Rabaca D., Rachlin E., Reich R., Ren Y., Reuter M., Richards S.,
Rives C., Rodriguez F., Rojas A., Ruiz S.J., Santa M., Sanders W.,
Santibanez J., Santos R., Savary G., Scherer S., Shen H., Shen Y.,
Sisson I., Sneed A., Sodergren E., Song X.-Z., Sorelle R.,
Svatek A., Taylor E., Taylor T., Thomas N., Thorn R., Thornton R.,
Trejos Z., Umami K., Vargo C., Verdusco D., Villasana D., Virk D.,
Volkov A., Waldron J., Walker B., Wang Q., Wang S., Warren J.,
Wei X., Wheeler D., Williams G., Williams R., Worley K., Wright R.,
Wu J., Yakub S., Yan K., Yaun Y., Yu F., Zhang J., Zhang L.,
Zhang Z., Zhou J., Weinstock G. and Gibbs R.

Direct Submission
Unpublished
2 (bases 1 to 169540)
Worley K.C.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help.tmc.edu
----- Project Information
Center project name: FBAB
Center clone name: CH240-22115
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 171082 bases at least Q40
Consensus quality: 172961 bases at least Q30
Consensus quality: 175246 bases at least Q20
Estimated insert size: 180839; sum-of-contigs estimation

```

```

-----
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 23 contigs. The true order of the pieces
is not known and their order in this sequence is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 3643: contig of 3643 bp in length
3644 3743: gap of unknown length
3744 13031: contig of 9288 bp in length
13032 13131: gap of unknown length
13132 16358: contig of 3127 bp in length
16359 24536: contig of unknown length
24537 24536: contig of 8078 bp in length
24537 27320: contig of unknown length
27321 27320: contig of 2784 bp in length
27321 30906: contig of 3486 bp in length
30907 39018: gap of unknown length
39019 39018: contig of 8012 bp in length
39019 43187: gap of unknown length
43188 43186: contig of 4068 bp in length
43187 43287: gap of unknown length
43288 54302: contig of 10916 bp in length
54303 54302: gap of unknown length
54303 56596: contig of 2294 bp in length
56597 56596: gap of unknown length
56597 59280: contig of 2584 bp in length
59281 59380: gap of unknown length
59381 64409: contig of 5029 bp in length
64410 64509: gap of unknown length
64510 78550: contig of 14041 bp in length
78551 78550: gap of unknown length
78551 81198: contig of 2548 bp in length
81199 81238: gap of unknown length
81239 86477: contig of 5179 bp in length
86478 86577: gap of unknown length
86578 101411: contig of 14834 bp in length
101412 101511: gap of unknown length
101512 104847: contig of 3336 bp in length
104848 104947: gap of unknown length
111324 111423: contig of 6376 bp in length
111424 125250: contig of 13827 bp in length
125251 125351: gap of unknown length
125352 135464: contig of 10113 bp in length
135465 150550: gap of unknown length
150551 150550: contig of 14987 bp in length
150551 157650: gap of unknown length
157651 157650: contig of 7000 bp in length
157651 169540: gap of unknown length
169540 169540: contig of 11790 bp in length.
-----
Location/Qualifiers
1. 169540
/organism=Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-22115"
3644. 3743
/estimated_length=unknown
13032. 13131
/estimated_length=unknown
16259. 16358
/estimated_length=unknown
24437. 24536
/estimated_length=unknown
27321. 27420
/estimated_length=unknown

```



```

GFFPQILGSRKRGKAPHRSPLS"
25102..25323
/gene="OSJNBA0073A21.3"
misc_feature
25102..25323
/gene="OSJNBA0073A21.3"
/notes="3' UTR transposable elements TNP2 protein-like"
30082..31019
/gene="OSJNBA0073A21.4"
misc_feature
30082..31019
/gene="OSJNBA0073A21.4"
/notes="probably inactive due to 5' exon missing in CDS
pseudogene, putative cyst nematode resistance protein"
pseudogene
complement(join(32626..32682,32727..32940,33036..33142))
/gene="OSJNBA0073A21.5"
misc_feature
complement(join(32626..32682,32727..32940,33036..33142))
/gene="OSJNBA0073A21.5"
/notes="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
complement(37925..38191)
/gene="OSJNBA0073A21.6"
misc_feature
complement(37925..38191)
/gene="OSJNBA0073A21.6"
/notes="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
complement(38706..41739)
/gene="OSJNBA0073A21.7"
misc_feature
complement(join(<38706..39166,39255..40218,40327..41444,
41538..>41739))
/notes="supported by full-length cDNA(s) : AKI19397"
complement(38706..41739)
/gene="OSJNBA0073A21.7"
misc_feature
/notes="contains full-length cDNA(s) : AKI19397
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
43643..49287
/gene="OSJNBA0073A21.8-1"
misc_feature
43643..49287
/gene="OSJNBA0073A21.8-2"
/notes="OSJNBA0073A21.8-2"
/notes="contains full-length cDNA(s) : AK066680
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
join(42848..42972,43128..43311,44727..44834,45075..45159,
45455..45690,45916..46089,46176..46223,46325..46384,
46481..46567,46673..46761,46875..46966,47776..47837,
47937..48086,48173..48388)
/gene="OSJNBA0073A21.8-1"
misc_feature
join(<42848..42972,43128..43311,44727..44834,45075..45159,
45455..45690,45916..46089,46176..46223,46325..46384,
46481..46567,46673..46761,46875..46966,47776..47837,
47937..48086,48173..48388)
/notes="OSJNBA0073A21.8-1"
/notes="contains EST(s) : AU173808(R3811),AU173809(R3811)
contains full-length cDNA(s) : AK062235"
/codon_start=1
/product="putative cytochrome P450"
/db_xref="GI:49388966"
/translation="MAITATAAAATPHWQADASPRHAACPALRGRRBLPVVRQ
SSGVDDPKSKRGGLDNASLLTNLSSGLSPKAPVAGVTDLGRPLFSLVDNWL
EBSYVYKLAFLPKAFVVSVDPIVARHLRLENACVYDGLVAILKPKGLKLEADL
TKQRKRVITPGFHALFDAMGVFTKCSERTIKLEELIERGEHGYKVTULLEAF
SNLADILIGLVGNFDFSDTKSPVTKAVYGTLPFAEHRSTFYIPYNNLELRTWIP
RQRKFHSLDKLVINDCLDSLKNKAKETRQEADVEKIQORDYSSLKDASLIRFEVDNRGA
DVEDQLRDDLMTLTIAGHETTAVALTWVSFLLAQNFSKMRKAAQAEVDSVLNETINV
DQKKLEYIRLIIIVEALLRLYPPPLLRRLRDKPLGGYNGAKEGVEIPACTGDIELS
IYNLHRSFYFWRDPEPEPRFSVPKDESIKGMAGDPDPSPCAMYPNEILADAFEL
PFGGPRKCVGDQFALLESTVALALLQKFDVLRGSPDEVNTGATHTHKSGLWCR
VRRRT"
43094..49322
/gene="OSJNBA0073A21.8-3"
misc_feature
join(<43094..43311,44727..44834,45075..45159,45455..45690,
45916..46089,46176..46223,46301..46384,46481..46567,
46673..46761,46875..46966,47776..47837,47937..48086,
48173..48412,48903..48922)
/notes="OSJNBA0073A21.8-3"
/notes="supported by full-length cDNA(s) : AKI00596"
43094..49322
/gene="OSJNBA0073A21.8-3"
misc_feature
/notes="contains full-length cDNA(s) : AKI00596
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
55652..56799
/gene="OSJNBA0073A21.9"
misc_feature
55652..56799
/gene="OSJNBA0073A21.9"
/notes="supported by full-length cDNA(s) : AK068418"
55672..56361
/gene="OSJNBA0073A21.9"
misc_feature
/notes="contains full-length cDNA(s) : AK068418"
/codon_start=1
/product="unknown protein"
Query Match 0.7%; Score 22; DB 4; Length 170051;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 986 AAGAAGAAATTGATAGATATTC 1007
DB 92973 AAGAAGAAATTGATAGATATTC 92952
RESULT 120
AC136070
LOCUS
DEFINITION
Rattus norvegicus clone CH230-321022, *** SEQUENCING IN PROGRESS
AC136070
AC136070.1 GI:24417941
VERSION
HTG; HTGS PHASE1
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus;
Muzny, D. Marie, Metzker, M. Lee, Abranzon, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Bunay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Devila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Deigado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebreyegorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, K.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

```

Karpthy,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorenuhewa,L., Louisseg,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mawhney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A.,
 Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
 Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
 Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okunolu,G.,
 Olarnpungagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,
 Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A.,
 Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E.,
 Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
 Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A.,
 Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savory,G., Scherer,S.,
 Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,
 Sison,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E.,
 Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A.,
 Svatek,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D.,
 Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,
 Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlaczky,R.,
 Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
 Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
 Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,
 Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 180005)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (29-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KDAP
 Center clone name: CH230-321022
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 141216 bases at least Q40
 Consensus quality: 146793 bases at least Q30
 Consensus quality: 151327 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 66 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1187: contig of 1187 bp in length
 * 1188 1287: gap of unknown length
 * 1288 2443: contig of 1156 bp in length
 * 2444 2544: gap of unknown length
 * 2544 3670: contig of 1127 bp in length
 * 3671 3770: gap of unknown length
 * 3771 4846: contig of 1076 bp in length
 * 4847 4946: gap of unknown length
 * 4947 6023: contig of 1076 bp in length
 * 6023 6122: gap of unknown length
 * 6122 7360: contig of 1238 bp in length
 * 7361 7460: gap of unknown length

7461 8749: contig of 1289 bp in length
 8750 8849: gap of unknown length
 8850 9996: contig of 1147 bp in length
 9997 10096: gap of unknown length
 10097 11833: contig of 1737 bp in length
 11834 13397: contig of 1464 bp in length
 13398 13497: gap of unknown length
 13498 15047: gap of unknown length
 15048 16173: contig of 1126 bp in length
 16174 18356: contig of 2083 bp in length
 18357 18457: gap of unknown length
 18458 20592: gap of unknown length
 20593 22782: contig of 2190 bp in length
 22783 23904: contig of 1022 bp in length
 23905 24004: gap of unknown length
 24005 26411: contig of 2407 bp in length
 26412 28511: gap of unknown length
 28512 28183: contig of 1572 bp in length
 28184 29417: contig of 1234 bp in length
 29418 31138: contig of 1621 bp in length
 31139 32338: gap of unknown length
 32339 32935: contig of 1597 bp in length
 32936 34816: contig of 1881 bp in length
 34817 36426: contig of 1510 bp in length
 36427 38231: contig of 1705 bp in length
 38232 38331: gap of unknown length
 38332 39365: contig of 1034 bp in length
 39366 41078: contig of 1613 bp in length
 41079 44036: contig of 2858 bp in length
 44037 46788: contig of 2652 bp in length
 46789 48521: contig of 1633 bp in length
 48522 50480: contig of 1859 bp in length
 50481 52769: contig of 2189 bp in length
 52770 55176: contig of 2307 bp in length
 55177 57071: contig of 1795 bp in length
 57072 59367: contig of 2196 bp in length
 59368 62414: contig of 2947 bp in length
 62415 65245: contig of 2731 bp in length
 65246 68118: contig of 2773 bp in length
 68119 71828: contig of 3610 bp in length
 71829 75096: contig of 3168 bp in length
 75097 77665: contig of 2469 bp in length
 77666 80433: contig of 2668 bp in length
 80434 83785: contig of 3252 bp in length
 83786 86640: contig of 2755 bp in length

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

```

* 86641 86740: gap of unknown length
* 86741 88907: contig of 2167 bp in length
* 88908 89007: gap of unknown length
* 89008 91758: contig of 2751 bp in length
* 91759 91858: gap of unknown length
* 91859 95513: contig of 3655 bp in length
* 95514 95613: gap of unknown length
* 95614 97638: contig of 1925 bp in length
* 97639 101107: gap of unknown length
* 101108 101207: contig of 3469 bp in length
* 101208 105010: gap of unknown length
* 105011 108251: contig of 3803 bp in length
* 108252 108351: gap of unknown length
* 108352 110710: contig of 3141 bp in length
* 110711 110810: gap of unknown length
* 110811 113807: contig of 2359 bp in length
* 113808 113907: gap of unknown length
* 113909 116303: contig of 2997 bp in length
* 116304 116403: gap of unknown length

Query Match      0.7%; Score 22; DB 12; Length 180005;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 492 AGAAATACACTGTTTATCT 513
Db 22158 AGAAATACACTGTTTATCT 22179

RESULT 121
AC174595/c
LOCUS AC174595
DEFINITION Bos taurus clone CH240-1507, WORKING DRAFT SEQUENCE, 15 unordered
pieces.
ACCESSION AC174595
VERSION AC174595.2 GI:86199977
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 187676)
Muzny,D., Adams,C., Agbai II,O., Allen,C., Alebrooks,S., Archer,P.,
Arredondo,H., Bandaranaike,D., Bangura,L., Beltan,B., Beltran,R.,
Beraducci,A., Biswal,J., Blyth,P., Bonham,H., Buhay,C., Burch,P.,
Cadoree,I., Canada,A., Cardenas,V., Carter,K., Cavazos,I.,
Chacko,J., Chahrouh,M., Chavez,D., Chen,A., Chen,G., Chen,R.,
Cheng,M.-T., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A.,
Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De
Anda,C., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H.,
Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K.,
Dziuda,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M.,
Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P.,
Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T.,
Gaskin,C., Gensch,S., Ghose,S., Gill,R., Gonzales,D.,
Gonzalez-Garay,M., Guevara,W., Holder,B., Hoiland,W., Haebleren,K.,
Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Havlak,P.,
Hawes,A., Hawkins,E., Hayes,S., Hemphill,L., Hernandez,J.,
Hines,S., Hitchens,M., Hodgson,A., Hogues,M., Hollins,B.,
Howell,L.T., Hulyk,S., Hume,J., Imo,K., Jackson,A., Jackson,L.,
Jacob,L., Jiang,H., Johnson,B., Johnson,R., Kalafus,K., Kelly,S.,
Keys,T., Khan,Z., King,L., Kovar,C., Kowis,A., Kowis,C., Lara,P.,
Leal,S., Lee,K., Lee,S., LeGall,F.I., Lemon,S., Lewis,L., Li,B.,
Li,Y., Li,Z., Linnell,M., Liu,W., Liu,Y.-S., Liu,Y., Liyanage,D.,
London,P., Lopez,J., Lorensuewa,L., Lozada,R., Luk,T., Madu,R.,
Maheshwari,M., Mahoney,C., Malloy,K., Mansouri,D., Martinez,E.,
McCllelland,H., McPherson,J., Mercadao,C., Metzker,M.,
Milosavljevic,A., Minja,E., Morgan,M., Morris,S., Munidaga,M.,
Murray,D., Nazarith,L., Ngo,D., Nguyen,N., Norwig-Eastaugh,E.,
Nott,A., Nwaokemeleah,O., Obregon,M., Ochi-Okorie,C., Odeh,E.,

```

Okwuonu,G., Okwuonu,K., Parker,D., Pasternak,S., Patel,B., Patel,V., Paul,H., Perez,A., Perez,L., Petrosino,J., Pham,T., Primus,E., Pu,L.-L., Puazo,M., Qin,X., Quinn,A., Quiroz,J., Rabata,D., Rachlin,E., Reigh,R., Ren,Y., Reuter,M., Richards,S., Rives,C., Rodriguez,F., Rojas,A., Ruiz,S.J., Sana,M., Sanders,W., Santibanez,J., Santos,R., Savery,G., Scherer,S., Shen,H., Shen,Y., Sisson,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Svatek,A., Taylor,E., Taylor,T., Thomas,N., Thorn,R., Thornton,R., Trejos,Z., Umami,K., Vargo,C., Verduzco,D., Villaseana,D., Virk,D., Volkov,A., Waldron,L., Walker,B., Wang,Q., Wang,S., Warren,J., Wei,X., Wheeler,D., Williams,G., Williams,R., Worley,K., Wright,R., Wu,J., Yakub,S., Yan,K., Yuan,Y., Yu,F., Zhang,J., Zhang,L., Zhang,Z., Zhou,J., Weinstein,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 187676)
Worley,K.C.
Direct Submission
Submitted (19-DEC-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 187676)

REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL
COMMENT

Bovine Genome Sequencing Consortium
Direct Submission
Submitted (31-JAN-2006) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jan 31, 2006 this sequence version replaced gi:83722266.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FATM
Center clone name: CH240-1507
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 184871 bases at least Q40
Consensus quality: 185703 bases at least Q30
Consensus quality: 186355 bases at least Q20
Estimated insert size: 189157; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a working draft sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
3726: contig of 3726 bp in length
3727 3776: gap of 50 bp
3777 17826: contig of 14050 bp in length
17827 17876: gap of 50 bp
17877 18992: contig of 2016 bp in length
18993 19942: gap of 50 bp

* 19943 46456: contig of 26514 bp in length
* 46457 46506: gap of 50 bp
* 46507 52968: contig of 6462 bp in length
* 52969 53018: gap of 50 bp
* 53019 53351: contig of 2333 bp in length
* 53352 55401: gap of 50 bp
* 55402 57888: contig of 2487 bp in length
* 57889 57938: gap of 50 bp
* 57939 90149: contig of 32211 bp in length
* 90150 90199: gap of 50 bp
* 90200 94796: contig of 4597 bp in length
* 94797 94846: gap of 50 bp
* 94847 101411: contig of 6565 bp in length
* 101412 101461: gap of 50 bp
* 101462 102671: contig of 1210 bp in length
* 102672 102721: gap of 50 bp
* 102722 124959: contig of 2238 bp in length
* 124960 125009: gap of 50 bp
* 125010 140365: contig of 15356 bp in length
* 140366 140415: gap of 50 bp
* 140416 185481: contig of 45066 bp in length
* 185482 185581: gap of unknown length
* 185582 187676: contig of 2095 bp in length.

FEATURES
source

Location/Qualifiers
1..187676
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-1507"
3727..3776
/estimated_length=50
17827..17876
/estimated_length=50
19893..19942
/estimated_length=50
46457..46506
/estimated_length=50
52969..53018
/estimated_length=50
55352..55401
/estimated_length=50
57889..57938
/estimated_length=50
90150..90199
/estimated_length=50
94797..94846
/estimated_length=50
101412..101461
/estimated_length=50
102672..102721
/estimated_length=50
124960..125009
/estimated_length=50
140366..140415
/estimated_length=50
185482..185581
/estimated_length=unknown

ORIGIN

Query Match 0.7%; Score 22; DB 12; Length 187676;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1416 TGATGTTATGGAAGGCAGAGA 1437
|||||
Db 67939 TGATGTTATGGAAGGCAGAGA 67918
|||||

RESULT 122
BX537287/c

LOCUS Zebrafish DNA sequence from clone CH211-20608 in linkage group 19,
DEFINITION complete sequence.
195505 bp DNA linear VRT 26-FEB-2004
AP000847/c

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BX537287 GI:42820909
HTG.
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 195505)
Barker,G.
Direct Submission
Submitted (26-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 25, 2004 this sequence version replaced gi:42592674.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfsh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep/Clone-derived
Zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhihong Bao and Sean Eddy, submitted), and those
beginning 'drr' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/projects/D_rerio/fishmask.shtml CH211-20608
is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.

FEATURES

Location/Qualifiers
1..195505
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-20608"
/clone_lib="CHORI-211"

ORIGIN

Query Match 0.7%; Score 22; DB 11; Length 195505;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3051 ATGCTCAAAATCAATGATACA 3072
|||||
Db 51574 ATGCTCAAAATCAATGATACA 51553
|||||

RESULT 123
AP000847/c

LOCUS Homo sapiens chromosome 11 clone RP11-773L2 map 11q24, WORKING
DEFINITION

```

ACCESSION      AP000847
VERSION        AP000847.2
KEYWORDS       HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE         Homo sapiens (human)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Homidae; Homo.
REFERENCE      1 (bases 1 to 196928)
               Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
               Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
               Homo sapiens 196,928 genomic DNA of 11q24
               Published Only in DataBase (1999)
REFERENCE      2 (bases 1 to 196928)
               Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
               Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
               Direct Submission
               Submitted (06-DEC-1999) Masahira Hattori, The Institute of Physical
               and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
               Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
               Japan (E-mail: hattori@gsc.riken.go.jp,
               URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923,
               Fax: 81-42-778-9924)
               On May 31, 2000 this sequence version replaced gi:6997679.
COMMENT       ----- Genome Center
               Center: RIKEN Genomic Sciences Center (GSC)
               Center code: RIKEN
               Web site: http://hgp.gsc.riken.go.jp/
               Contact: hattori@gsc.riken.go.jp
               ----- Project Information
               Center project name: HumDraft11
               Center clone name: Rpl11-773L2
               ----- Summary Statistics
               Sequencing vector: PCR products; 100% of reads
               Chemistry: Dye-terminator ET-amersham; 100% of reads
               Assembly program: Phrap; version 0.990329
               Consensus quality: 179745 bases at least Q40
               Consensus quality: 180443 bases at least Q30
               Consensus quality: 191783 bases at least Q20
               Insert size: 193928; sum-of-contigs
               Quality coverage: 4.17X in Q20 bases; sum-of-contigs
               -----
NOTE: This is a 'working draft' sequence. It currently consists of
31 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1
13348 contig of 13348 bp in length
13449 contig of 100 bp
26989 contig of 11607 bp in length
27090 contig of 12107 bp in length
51004 contig of 9975 bp in length
51009 contig of 9901 bp in length
82905 contig of 9963 bp in length
82906 contig of 100 bp
102869 contig of 100 bp
102869 contig of 100 bp
110834 contig of 7866 bp in length
110835 contig of 100 bp
110935 contig of 100 bp
118975 contig of 7940 bp in length
118975 contig of 100 bp
118975 contig of 100 bp
126648 contig of 7673 bp in length
126648 contig of 100 bp
135519 contig of 8771 bp in length
135519 contig of 100 bp
141552 contig of 5833 bp in length
141552 contig of 100 bp
148610 contig of 7058 bp in length
148610 contig of 100 bp
154319 contig of 5610 bp in length
154320 contig of 100 bp
160564 contig of 6144 bp in length
160564 contig of 100 bp
167021 contig of 6357 bp in length
167021 contig of 100 bp
171317 contig of 4096 bp in length
171317 contig of 100 bp
173789 contig of 4472 bp in length
173789 contig of 100 bp
178186 contig of 2297 bp in length
178186 contig of 100 bp
180624 contig of 2338 bp in length
180624 contig of 100 bp
180724 contig of 1630 bp in length
182354 contig of 100 bp
182354 contig of 100 bp
184552 contig of 2168 bp in length
184552 contig of 100 bp
184722 contig of 100 bp
184722 contig of 100 bp
186904 contig of 2082 bp in length
186904 contig of 100 bp
189123 contig of 2219 bp in length
189123 contig of 100 bp
189223 contig of 1934 bp in length
191157 contig of 100 bp
191256 contig of 100 bp
192980 contig of 1724 bp in length
192980 contig of 100 bp
192981 contig of 100 bp
193080: gap of 100 bp

```

Sequence updated (26-May-2000)

* NOTE: This is a 'working draft' sequence. It currently consists of 31 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

13348: contig of 13348 bp in length
13449: gap of 100 bp
26989: contig of 11607 bp in length
27090: contig of 12107 bp in length
51004: contig of 9975 bp in length
51009: contig of 9901 bp in length
82905: contig of 9963 bp in length
82906: contig of 100 bp
102869: contig of 100 bp
102869: contig of 100 bp
110834: contig of 7866 bp in length
110835: contig of 100 bp
110935: contig of 100 bp
118975: contig of 7940 bp in length
118975: contig of 100 bp
118975: contig of 100 bp
126648: contig of 7673 bp in length
126648: contig of 100 bp
135519: contig of 8771 bp in length
135519: contig of 100 bp
141552: contig of 5833 bp in length
141552: contig of 100 bp
148610: contig of 7058 bp in length
148610: contig of 100 bp
154319: contig of 5610 bp in length
154320: contig of 100 bp
160564: contig of 6144 bp in length
160564: contig of 100 bp
167021: contig of 6357 bp in length
167021: contig of 100 bp
171317: contig of 4096 bp in length
171317: contig of 100 bp
173789: contig of 4472 bp in length
173789: contig of 100 bp
178186: contig of 2297 bp in length
178186: contig of 100 bp
180624: contig of 2338 bp in length
180624: contig of 100 bp
180724: contig of 1630 bp in length
182354: contig of 100 bp
182354: contig of 100 bp
184552: contig of 2168 bp in length
184552: contig of 100 bp
184722: contig of 100 bp
184722: contig of 100 bp
186904: contig of 2082 bp in length
186904: contig of 100 bp
189123: contig of 2219 bp in length
189123: contig of 100 bp
189223: contig of 1934 bp in length
191157: contig of 100 bp
191256: contig of 100 bp
192980: contig of 1724 bp in length
192980: contig of 100 bp
192981: contig of 100 bp
193080: gap of 100 bp


```
repeat_region /clone="RP23-354D22"
repeat_region 935..1281 /rpt_family="MaLR"
repeat_region 1296..1508 /rpt_family="MaLR"
repeat_region 1536..1718 /rpt_family="MaLR"
repeat_region 1870..2093 /rpt_family="MaLR"
repeat_region 1870..2093 /rpt_family="MaLR"
repeat_region 2835..3197 /rpt_family="ERV1"
repeat_region 4304..4781 /rpt_family="MaLR"
repeat_region 4304..4781 /rpt_family="MaLR"
repeat_region 6763..6882 /rpt_family="B4"
repeat_region 6953..7247 /rpt_family="MIR"
repeat_region 10130..10214 /rpt_family="B4"
repeat_region 10626..11096 /rpt_family="B4"
repeat_region 11115..11173 /rpt_family="RMER19B"
repeat_region 11319..11557 /rpt_family="RMER19A"
repeat_region 13777..13950 /rpt_family="MaLR"
misc_feature /note="Sequence derived from PCR product of project DNA."
repeat_region 14130..14374 /rpt_family="MaLR"
repeat_region 14380..14603 /rpt_family="MaLR"
repeat_region 14769..14965 /rpt_family="MaLR"
repeat_region 15019..15835 /rpt_family="MaLR"
repeat_region 15803..16231 /rpt_family="MaLR"
repeat_region 16235..16601 /rpt_family="MaLR"
repeat_region 19268..19434 /rpt_family="MaLR"
repeat_region 19847..20011 /rpt_family="B2"
repeat_region 22856..23011 /rpt_family="L1"
repeat_region 24412..24612 /rpt_family="B4"
repeat_region 24740..24780 /rpt_family="RMER10B"
repeat_region 24975..25195 /rpt_family="RMER17C"
repeat_region 26153..26442 /rpt_family="B4"
repeat_region 26878..27099 /rpt_family="L1"
repeat_region 28861..28997 /rpt_family="MER1_type"
repeat_region 28998..29146 /rpt_family="Alu"
repeat_region 30673..30827 /rpt_family="ERVK"
repeat_region 31518..31628 /rpt_family="MaLR"
repeat_region 34256..34311 /rpt_family="U3"
repeat_region 35039..35245 /rpt_family="L1"
repeat_region 35249..35405 /rpt_family="L1"
repeat_region 35646..35711 /rpt_family="L1"
repeat_region 36070..36505 /rpt_family="L1"
```

```
repeat_region /rpt_family="MaLR"
repeat_region 36798..37017 /rpt_family="B4"
repeat_region 37997..38691 /rpt_family="L1"
repeat_region 40345..40407 /rpt_family="ERV1"
repeat_region 41048..41202 /rpt_family="L1"
repeat_region 41899..42336 /rpt_family="L1"
repeat_region 42787..42842 /rpt_family="ID"
misc_feature 42949..43062 /note="Sequence derived from PCR product of project DNA."
repeat_region 43183..43298 /rpt_family="L1"
repeat_region 43745..43854 /rpt_family="MIR"
repeat_region 44932..44983 /rpt_family="ERV1"
repeat_region 44963..45013 /rpt_family="ERV1"
repeat_region 45084..45134 /rpt_family="ERV1"
repeat_region 45191..45240 /rpt_family="ERV1"
repeat_region 45249..45315 /rpt_family="L1"
repeat_region 45321..45375 /rpt_family="ERV1"
repeat_region 45355..45406 /rpt_family="ERV1"
repeat_region 47156..47391 /rpt_family="L1"
repeat_region 49636..49840 /rpt_family="B4"
repeat_region 51112..51333 /rpt_family="MER1_type"
repeat_region 55474..55547 /rpt_family="Alu"
repeat_region 57783..57978 /rpt_family="B2"
repeat_region 58531..58624 /rpt_family="B2"
```

Query Match 0.7%; Score 22; DB 6; Length 207817;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2064 AGAAATTTTCCTTTGAAAGT 2085
Db 83187 AGAAATTTTCCTTTGAAAGT 83166

RESULT 125
AC151257/c
LOCUS AC151257
DEFINITION Bos taurus clone CH240-342N10, WORKING DRAFT SEQUENCE, 17 unordered
pieces.
ACCESSION AC151257
VERSION AC151257.3 GI:68285941
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 210651)
Muzny,D.Marte., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayoyagi,M., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FBHW
Center clone name: CH240-342N10
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 201631 bases at least Q40
Consensus quality: 203740 bases at least Q30
Consensus quality: 205693 bases at least Q20
Estimated insert size: 210305; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2052: contig of 2052 bp in length
2053: gap of 1263 bp
3315: contig of 2995 bp in length
6310: gap of unknown length
6410: contig of 6459 bp in length
12869: contig of 50 bp
12919: gap of 50 bp
12970: contig of 22354 bp in length
35273: gap of 50 bp
35324: contig of 25066 bp in length
60339: gap of 50 bp
60390: contig of 14194 bp in length
60440: gap of 50 bp
74633: contig of 26899 bp in length
74683: gap of 50 bp
74684: contig of 26634 bp in length
101582: gap of 50 bp
101632: contig of 15338 bp in length
128266: gap of 50 bp
128316: contig of 47008 bp in length
128371: gap of 50 bp
143704: contig of 7558 bp in length
143705: gap of 50 bp
190713: contig of 4686 bp in length
190763: gap of 50 bp
198321: contig of 1525 bp in length
198371: gap of 50 bp
203057: contig of 1478 bp in length
203107: gap of 50 bp
204631: contig of unknown length
204682: gap of 1374 bp in length
206160: contig of unknown length
206260: gap of 1787 bp in length
207634: contig of unknown length
207733: gap of 50 bp
209520: gap of unknown length
209521: contig of 1031 bp in length.
209621: Location/Qualifiers
1..210651
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-342N10"
2053..3315
/estimated_length=1263
6311..6410
/estimated_length=unknown
12870..12919
/estimated_length=50
35274..35323
/estimated_length=50
60390..60439

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FBHW
Center clone name: CH240-342N10
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 201631 bases at least Q40
Consensus quality: 203740 bases at least Q30
Consensus quality: 205693 bases at least Q20
Estimated insert size: 210305; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2052: contig of 2052 bp in length
2053: gap of 1263 bp
3315: contig of 2995 bp in length
6310: gap of unknown length
6410: contig of 6459 bp in length
12869: contig of 50 bp
12919: gap of 50 bp
12970: contig of 22354 bp in length
35273: gap of 50 bp
35324: contig of 25066 bp in length
60339: gap of 50 bp
60390: contig of 14194 bp in length
60440: gap of 50 bp
74633: contig of 26899 bp in length
74683: gap of 50 bp
74684: contig of 26634 bp in length
101582: gap of 50 bp
101632: contig of 15338 bp in length
128266: gap of 50 bp
128316: contig of 47008 bp in length
128371: gap of 50 bp
143704: contig of 7558 bp in length
143705: gap of 50 bp
190713: contig of 4686 bp in length
190763: gap of 50 bp
198321: contig of 1525 bp in length
198371: gap of 50 bp
203057: contig of 1478 bp in length
203107: gap of 50 bp
204631: contig of unknown length
204682: gap of 1374 bp in length
206160: contig of unknown length
206260: gap of 1787 bp in length
207634: contig of unknown length
207733: gap of 50 bp
209520: gap of unknown length
209521: contig of 1031 bp in length.
209621: Location/Qualifiers
1..210651
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-342N10"
2053..3315
/estimated_length=1263
6311..6410
/estimated_length=unknown
12870..12919
/estimated_length=50
35274..35323
/estimated_length=50
60390..60439

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FBHW
Center clone name: CH240-342N10
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 201631 bases at least Q40
Consensus quality: 203740 bases at least Q30
Consensus quality: 205693 bases at least Q20
Estimated insert size: 210305; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2052: contig of 2052 bp in length
2053: gap of 1263 bp
3315: contig of 2995 bp in length
6310: gap of unknown length
6410: contig of 6459 bp in length
12869: contig of 50 bp
12919: gap of 50 bp
12970: contig of 22354 bp in length
35273: gap of 50 bp
35324: contig of 25066 bp in length
60339: gap of 50 bp
60390: contig of 14194 bp in length
60440: gap of 50 bp
74633: contig of 26899 bp in length
74683: gap of 50 bp
74684: contig of 26634 bp in length
101582: gap of 50 bp
101632: contig of 15338 bp in length
128266: gap of 50 bp
128316: contig of 47008 bp in length
128371: gap of 50 bp
143704: contig of 7558 bp in length
143705: gap of 50 bp
190713: contig of 4686 bp in length
190763: gap of 50 bp
198321: contig of 1525 bp in length
198371: gap of 50 bp
203057: contig of 1478 bp in length
203107: gap of 50 bp
204631: contig of unknown length
204682: gap of 1374 bp in length
206160: contig of unknown length
206260: gap of 1787 bp in length
207634: contig of unknown length
207733: gap of 50 bp
209520: gap of unknown length
209521: contig of 1031 bp in length.
209621: Location/Qualifiers
1..210651
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-342N10"
2053..3315
/estimated_length=1263
6311..6410
/estimated_length=unknown
12870..12919
/estimated_length=50
35274..35323
/estimated_length=50
60390..60439

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FBHW
Center clone name: CH240-342N10
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 201631 bases at least Q40
Consensus quality: 203740 bases at least Q30
Consensus quality: 205693 bases at least Q20
Estimated insert size: 210305; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2052: contig of 2052 bp in length
2053: gap of 1263 bp
3315: contig of 2995 bp in length
6310: gap of unknown length
6410: contig of 6459 bp in length
12869: contig of 50 bp
12919: gap of 50 bp
12970: contig of 22354 bp in length
35273: gap of 50 bp
35324: contig of 25066 bp in length
60339: gap of 50 bp
60390: contig of 14194 bp in length
60440: gap of 50 bp
74633: contig of 26899 bp in length
74683: gap of 50 bp
74684: contig of 26634 bp in length
101582: gap of 50 bp
101632: contig of 15338 bp in length
128266: gap of 50 bp
128316: contig of 47008 bp in length
128371: gap of 50 bp
143704: contig of 7558 bp in length
143705: gap of 50 bp
190713: contig of 4686 bp in length
190763: gap of 50 bp
198321: contig of 1525 bp in length
198371: gap of 50 bp
203057: contig of 1478 bp in length
203107: gap of 50 bp
204631: contig of unknown length
204682: gap of 1374 bp in length
206160: contig of unknown length
206260: gap of 1787 bp in length
207634: contig of unknown length
207733: gap of 50 bp
209520: gap of unknown length
209521: contig of 1031 bp in length.
209621: Location/Qualifiers
1..210651
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-342N10"
2053..3315
/estimated_length=1263
6311..6410
/estimated_length=unknown
12870..12919
/estimated_length=50
35274..35323
/estimated_length=50
60390..60439

```

/estimated_length=50
74634 .74683
/estimated_length=50
101583 .101632
/estimated_length=50
128267 .128316
/estimated_length=50
143655 .143704
/estimated_length=50
190713 .190762
/estimated_length=50
198321 .198370
/estimated_length=50
203057 .203106
/estimated_length=50
204632 .204681
/estimated_length=50
206160 .206259
/estimated_length=unknown
207634 .207733
/estimated_length=unknown
209521 .209620
/estimated_length=unknown

ORIGIN
Query Match      0.7%  Score 22;  DB 12;  Length 210651;
Best Local Similarity 100.0%;  Pred.No. 5.8;  0;  Indels  0;  Gaps  0;
Matches 22;  Conservative  0;  Mismatches  0;

QY  1506 AAATATCCATGACATCTTTCAT 1527
Db  53204 AAATATCCATGACATCTTTCAT 53183
|||||
|||||

RESULT 126
AC167594
LOCUS
DEFINITION
AC167594
AC167594.2 GI:85665454
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Bos taurus
Bos taurus (cattle)

Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 210788)
Muzny, D., Adams, C., Agbai II, O., Allen, C., Albrooks, S., Archer, P.,
Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,
Beraducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
Chacko, J., Chahrouh, M., Chavez, D., Chen, A., Chen, G., Chen, R.,
Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Carroll, L., De
Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H.,
Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
Dziulida, D., Egan, A., Escoto, M., Espinosa, V., Eugene, C., Fa, M.,
Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,
Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,
Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D.,
Gonzalez-Garay, M., Guevara, W., Holder, M., Hoiland, W., Haeblerlen, K.,
Hall, B., Hamid, H., Hamilton, K., Harber, B., Harris, R., Havlak, P.,
Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,
Hines, S., Hitchens, M., Hodgson, A., Hogue, M., Hollins, B.,
Howell, L. T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,
Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kallatus, K., Kelly, S.,
Key, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F.,
Lea, S., Lee, K., Lee, S., LeGall, F. I., Lemon, S., Lewis, L., Li, B.,
Li, X., Li, Z., Limell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D.,
London, F., Lopez, J., Lorensuewa, L., Lozada, R., Luk, T., Madu, R.,
Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
McClelland, H., McPherson, J., Mercadao, C., Metzker, M.,

```

Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M., Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Estaugh, E., Nott, A., Nwaokemele, O., Obregon, M., Ochi-Okorie, C., Odeh, E., Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J., Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S., Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sheed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorpe, R., Thornton, R., Trejos, Z., Usmani, K., Varco, C., Verduzco, D., Villaseca, D., York, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstein, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 210788)
Worley, K. C.
Direct Submission
Submitted (01-SEP-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 210788)

Bovine Genome Sequencing Consortium
Direct Submission
Submitted (23-JAN-2006) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jan 23, 2006 this sequence version replaced gi:74039134.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the ends of the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc.help@bcm.tmc.edu
----- Project Information
Center Project name: FLPI1
Center Clone name: CH240-137M9
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 199274 bases at least Q40
Consensus quality: 202130 bases at least Q30
Consensus quality: 204587 bases at least Q20
Estimated insert size: 210018; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1
* 23387: contig of 23387 bp in length
* 23388 23437: gap of 50 bp
* 23438 36183: contig of 12746 bp in length

* 36184 36233: gap of 50 bp
* 36234 38005: contig of 1772 bp in length
* 38006 38055: gap of 50 bp
* 38056 51012: contig of 12957 bp in length
* 51013 73800: contig of 225 bp
* 51238 73800: contig of 22563 bp in length
* 73801 73927: gap of 127 bp
* 73928 78791: contig of 4864 bp in length
* 78792 78841: gap of 50 bp
* 78842 87408: contig of 8567 bp in length
* 87409 87458: gap of 50 bp
* 87459 96969: contig of 9511 bp in length
* 96970 97019: gap of 50 bp
* 97020 125105: contig of 28086 bp in length
* 125106 125155: gap of 50 bp
* 125156 132888: contig of 7733 bp in length
* 132889 132938: gap of 50 bp
* 132939 138261: contig of 5323 bp in length
* 138262 138311: gap of 50 bp
* 138312 163617: contig of 25306 bp in length
* 163618 163667: gap of 50 bp
* 163668 167075: contig of 3408 bp in length
* 167076 167176: gap of unknown length
* 167177 168970: contig of 1795 bp in length
* 168971 169310: gap of 340 bp
* 169311 175771: contig of 6461 bp in length
* 175772 175821: gap of 50 bp
* 175822 185353: contig of 9532 bp in length
* 185354 185496: gap of 143 bp
* 185497 186995: contig of 1499 bp in length
* 186996 187095: gap of unknown length
* 187096 197305: contig of 10210 bp in length
* 197306 197355: gap of 50 bp
* 197356 199414: contig of 2059 bp in length
* 199415 199514: gap of unknown length
* 200526 200526: contig of 1012 bp in length
* 200527 200626: gap of unknown length
* 200627 201710: contig of 1084 bp in length
* 201711 201810: gap of unknown length
* 201811 203807: contig of 1997 bp in length
* 203808 203907: gap of unknown length
* 203908 205769: contig of 1862 bp in length
* 205770 205869: gap of unknown length
* 205870 207636: contig of 1767 bp in length
* 207637 207736: gap of unknown length
* 207737 210788: contig of 3052 bp in length.

FEATURES
source
Location/Qualifiers
1. .210788
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-137M9"
23388. .23437
/estimated_length=50
36184. .36233
/estimated_length=50
38006. .38055
/estimated_length=50
51013. .51237
/estimated_length=225
73801. .73927
/estimated_length=127
78792. .78841
/estimated_length=50
87409. .87458
/estimated_length=50
96970. .97019
/estimated_length=50
125106. .125155
/estimated_length=50
132889. .132938
/estimated_length=50
138262. .138311

/estimated_length=50
163618. .163667
/estimated_length=50
167076. .167175
/estimated_length=unknown
168971. .169310
/estimated_length=340
175772. .175821
/estimated_length=50
185354. .185496
/estimated_length=143
186996. .187095
/estimated_length=unknown
197306. .197355
/estimated_length=50
199415. .199514
/estimated_length=unknown

Query Match
Best Local Similarity 0.7%; Score 22; DB 12; Length 210788;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1068 TCTATATGAGAGAAATGATGAA 1089
|||||
DB 37310 TCTATATGAGAGAAATGATGAA 37331

RESULT 127
AC162627/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC162627 214864 bp DNA linear HTG 01-JUL-2005
Bos taurus clone CH240-124D18, *** SEQUENCING IN PROGRESS ***
unordered pieces.
AC162627 GI:68301357
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Bos taurus (cattle)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 214864)
Muzny,D,Marie., Metzker,M,Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alabrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Blyth,P., Brown,M.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,E.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geert,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpthy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokemele,O., Okwuonu,G., Olarnpunaagoon,A., Pal,S., Parks,K.,

Query Match 0.7%; Score 22; DB 12; Length 214864;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1416 TGATGTTATGGAAGCAGAGA 1437
 |||||
 Db 158460 TGATGTTATGGAAGCAGAGA 158439

RESULT 128
 AC156845 222380 bp DNA linear HTG 01-JUL-2005
 LOCUS Bos taurus clone CH240-65E3, *** SEQUENCING IN PROGRESS ***, 23
 DEFINITION unordered pieces.

ACCESSION AC156845.2 GI:68303004
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS Bos taurus (cattle)
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 222380)
 Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, R., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gregoiregis, E., Geer, K., Gill, R., Grady, M., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensuhewa, L., Louissegh, H., Lozado, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Mlosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstock, G. and Gibbs, R. A.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Direct Submission
 Unpublished
 2 (bases 1 to 222380)
 Worley, K. C.
 Direct Submission
 Submitted (05-FEB-2005) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 222380)
 Cow Genome Sequencing Consortium.
 Direct Submission
 Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jun 29, 2005 this sequence version replaced gi:58652202.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: FDTG
 Center clone name: CH240-65E3
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 207604 bases at least Q40
 Consensus quality: 210552 bases at least Q30
 Consensus quality: 213132 bases at least Q20
 Estimated insert size: 214079; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * This is a 'working draft' sequence. It currently
 * consists of 23 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

Accession	Contig	Length
1	1539: contig of 1539 bp in length	1539
1540	1871: gap of 332 bp	1540
1872	7640: contig of 5769 bp in length	1872
7641	7874: gap of 234 bp	7641
7875	23979: contig of 16105 bp in length	7875
23980	24029: gap of 50 bp	23980
24030	29654: contig of 5625 bp in length	24030
29655	29865: gap of 211 bp	29655
29866	34956: contig of 5091 bp in length	29866
34957	35006: gap of 50 bp	34957
35007	41109: contig of 6103 bp in length	35007
41110	41688: gap of 579 bp	41110
41689	48375: contig of 6687 bp in length	41689
48376	48425: gap of 50 bp	48376
48426	84184: contig of 35759 bp in length	48426
84185	84234: gap of 50 bp	84185
84235	88842: contig of 4608 bp in length	84235
88843	88892: gap of 50 bp	88843
88893	98098: contig of 9206 bp in length	88893
98099	98148: gap of 50 bp	98099

Db	211020	TCGAAAGGCAGAGACTCATTGA	211041
RESULT 129			
AC105304			
LOCUS			
DEFINITION	AC105304	229816 bp	DNA linear ROD 16-SEP-2003
	Mus musculus strain C57BL/6J clone rp23-71k5	map 17, complete	sequence.
ACCESSION	AC105304	GI:25815339	
VERSION	AC105304		
KEYWORDS	HTG		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;		
	Sciurognathi; Muridea; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 229816)		
AUTHORS	Deschamps,S., Li,Y., Hu,X., Brathwaite,M., Waeltz,P., Nagaraja,R.		
	and Roe,B.A.		
TITLE	Mus musculus BAC Clone rp23-71k5		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 229816)		
AUTHORS	Hu,X. and Roe,B.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-DEC-2001) Department Of Chemistry And Biochemistry,		
	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,		
	OK 73019, USA		
REFERENCE	3 (bases 1 to 229816)		
AUTHORS	Deschamps,S., Li,Y., Hu,X. and Roe,B.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-NOV-2002) Department Of Chemistry And Biochemistry,		
	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,		
	OK 73019, USA		
REFERENCE	4 (bases 1 to 229816)		
AUTHORS	Deschamps,S., Li,Y., Hu,X. and Roe,B.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-NOV-2002) Department Of Chemistry And Biochemistry,		
	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,		
	OK 73019, USA		
REFERENCE	5 (bases 1 to 229816)		
AUTHORS	Deschamps,S., Li,Y., Hu,X., Brathwaite,M., Waeltz,P., Nagaraja,R.		
	and Roe,B.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-MAR-2003) Department Of Chemistry And Biochemistry,		
	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,		
	OK 73019, USA		
REFERENCE	6 (bases 1 to 229816)		
AUTHORS	Deschamps,S., Li,Y., Hu,X., Brathwaite,M., Waeltz,P., Nagaraja,R.		
	and Roe,B.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-SEP-2003) Department Of Chemistry And Biochemistry,		
	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,		
	OK 73019, USA		
COMMENT	On Nov 28, 2002 this sequence version replaced gi:25141046.		
	----- Genome Center		
	Center: Department Of Chemistry And Biochemistry		
	The University Of Oklahoma		
	Center code:UOKNOR		

FEATURES	Location/Qualifiers		
source	1. .229816		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9913"		
	/clone="CH240-65E3"		
gap	1540. .1871		
	/estimated_length=332		
gap	7641. .7874		
	/estimated_length=234		
gap	23980. .24029		
	/estimated_length=50		
gap	29655. .29865		
	/estimated_length=211		
gap	34957. .35006		
	/estimated_length=50		
gap	4110. .41688		
	/estimated_length=579		
gap	48376. .48425		
	/estimated_length=50		
gap	84185. .84234		
	/estimated_length=50		
gap	88843. .88892		
	/estimated_length=50		
gap	98099. .98148		
	/estimated_length=50		
gap	101063. .102005		
	/estimated_length=943		
gap	181091. .182305		
	/estimated_length=1215		
gap	184196. .186080		
	/estimated_length=1885		
gap	187631. .187680		
	/estimated_length=50		
gap	202319. .202368		
	/estimated_length=50		
gap	205028. .205077		
	/estimated_length=50		
gap	209277. .209326		
	/estimated_length=50		
gap	215615. .215714		
Query Match	0.7%; Score 22; DB 12; Length 222380;		
Best Local Similarity	100.0%; Pred. No. 5.8;		
Matches	22; Conservative	0; Mismatches	0; Indels
		0; Gaps	0;
QY	1424	TCGAAAGGCAGAGACTCATTGA	1445

QY	2064	AGAAATTTTCTCTTTTGAAAGT	2085
DB	43236	AGAAATTTTCTCTTTTGAAAGT	43257
RESULT	130		
AC097286/C			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
AUTHORS			
REFERENCE			
Direct Submission			
Submitted (14-OCT-2001)			
of Molecular and Human Genetics,			
Baylor College of Medicine, One			
Baylor Plaza, Houston, TX 77030, USA			
3 (bases 1 to 231295)			
Rat Genome Sequencing Consortium.			
Direct Submission			
Submitted (10-MAY-2003)			
Human Genome Sequencing Center, Department			
of Molecular and Human Genetics, Baylor College of Medicine, One			
Baylor Plaza, Houston, TX 77030, USA			
On May 10, 2003 this sequence version replaced gi:24635699.			
The sequence in this assembly is a combination of BAC based reads			
and whole genome shotgun sequencing reads assembled using Atlas			
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described			
in the feature table below represents a scaffold in the Atlas			
assembly (a 'contig-scaffold'). Within each contig-scaffold,			
individual sequence contigs are ordered and oriented, and separated			
by sized gaps filled with Ns to the estimated size. The sequence			
may extend beyond the ends of the clone and there may be sequence			
contigs within a contig-scaffold that consist entirely of whole			
genome shotgun sequence reads. Both end sequences and whole genome			
shotgun sequence only contigs will be indicated in the feature			
table.			
----- Genome Center			
Center: Baylor College of Medicine			
Center code: BCM			
Web site: http://www.hgsc.bcm.tmc.edu/			
Contact: hgsc-help@bcm.tmc.edu			
----- Project Information			
Center project name: GGOZ			
Center Clone name: CH230-19G10			
----- Summary Statistics			
Assembly program: Atlas 3.0;			
Consensus quality: 223363 bases at least Q40			
Consensus quality: 225310 bases at least Q30			
Consensus quality: 226539 bases at least Q20			
Estimated insert size: 231185; sum-of-contigs estimation			
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation			

* NOTE: Estimated insert size may differ from sequence length			
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)			
* NOTE: This is a 'working draft' sequence. It currently			
consists of 3 contigs. The true order of the pieces			
is not known and their order in this sequence record is			
arbitrary. Gaps between the contigs are represented as			
runs of N, but the exact sizes of the gaps are unknown.			
* This record will be updated with the finished sequence.			
* As soon as it is available and the accession number will			
be preserved.			
*			
1 228572: contig of 228572 bp in length			
*			
228573 228672: gap of unknown length			
*			
228673 230105: contig of 1433 bp in length			
*			
230106 230205: gap of unknown length			
*			
230206 231295: contig of 1090 bp in length.			
Location/Qualifiers			
1..231295			
/organism="Rattus norvegicus"			
/mol_type="genomic DNA"			
/db_xref="taxon:10116"			
/clones="CH230-19G10"			
1421..2288			
/note="clone boundary			
clone end:Sp5			
site:EcoRI			
end sequence:BH269139"			
complement(226435..231295)			
/note="clone boundary			
clone end:T7			
site:EcoRI			
end sequence:BH269138"			
228573..228672			
/estimated_length=unknown			
230106..230205			
FEATURES			
Source			
misc_feature			
misc_feature			
gap			
gap			


```

L. .294090
/organism="Rattus norvegicus"
/mol_type="genomic DNA"

```

```

/db_xref="taxon:10116"
/clone="CH230-204F23"
1..1665
/notes="ags end extension
clone_end:sp6"
3110..4758
/notes="wgs_end extension
clone_end:sp6"
5478..6171
/notes="clone boundary
clone_end:sp6
site:EcoRI
end_sequence:RWBWJ36TV"
complement(288029..288819)
/notes="clone boundary
clone_end:T7
site:EcoRI
end_sequence:RWBWJ36TJ"

ORIGIN
Query Match 0.7%; Score 22; DB 12; Length 294090;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1853 TGACAGGCTGACTGACCGTGG 1874
|||||
Db 268679 TGACAGGCTGACTGACCGTGG 268700

RESULT 133
CQ690820
LOCUS 282 bp DNA
DEFINITION Sequence 35746 from Patent WO02070737.
ACCESSION CQ690820
VERSION CQ690820.1 GI:42228399
KEYWORDS Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE
1 Liew,C.C., Marshall,W.E. and Zhang,H.
AUTHORS Compositions and methods relating to osteoarthritis
TITLE Patent: WO 02070737-A 35746 12-SEP-2002;
JOURNAL Chondrogene Inc. (CA)
FEATURES
source
1..282
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2422 GGCTGTCCTATGGAGGATAC 2442
|||||
Db 233 GGCTGTCCTATGGAGGATAC 253

RESULT 134
AJ865196
LOCUS 400 bp DNA linear
DEFINITION Cocos nucifera microsatellite DNA, clone CnCir 34.
ACCESSION AJ865196
VERSION AJ865196.1 GI:57898801
KEYWORDS microsatellite; repetitive DNA.
SOURCE Cocos nucifera (coconut palm)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Arecales; Arecoidae;

/misc_feature misc_feature misc_feature misc_feature misc_feature
Cocoseae; Butiinae; Cocos.
1 Lebrun,P.
AUTHORS CRDXRIT: A genetic coconut map
TITLE Unpublished
REFERENCE 2 (bases 1 to 400)
AUTHORS Lebrun,P.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-2004) Lebrun P., CIRAD, TA 80/03 Avenue du Val de
Montferand, 34398 Montpellier Cedex 5, FRANCE
FEATURES
source
1..400
Location/Qualifiers
/organism="Cocos nucifera"
/mol_type="genomic DNA"
/db_xref="taxon:13894"
/clone="CnCir_34"
/tissue_type="leaf"
/country="Malaysia"
96..101
/notes="Forward Primer: CCTCCCACTCCGTATT"
168..201
/notes="microsatellite"
/rpt_type="RANDEM"
/rpt_unit_range=168..201
complement(369..383)
/notes="Reverse Primer: GGCACGGCAAGTAG"

ORIGIN
Query Match 0.7%; Score 21; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3100 AAGAAATTACTAAAAA 3120
|||||
Db 29 AAGAAATTACTAAAAA 49

RESULT 135
AR280920
LOCUS 535 bp DNA linear
DEFINITION Sequence 428 from patent US 6518237.
ACCESSION AR280920
VERSION AR280920.1 GI:29716390
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 535)
AUTHORS Yugui,J Dillon,D.C.; Mitcham,J.L. and Xu,J.
TITLE Compositions for treatment and diagnosis of breast cancer and
methods for their use
JOURNAL Patent: US 6518237-A 428 11-FEB-2003;
FEATURES
source
1..535
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2422 GGCTGGTCCTATGGAGGATAC 2442
|||||
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 136
AR283416
LOCUS 535 bp DNA linear
DEFINITION Sequence 428 from patent US 6528054.
ACCESSION AR283416
```

Thu Jun 22 09:04:25 2006

```
JOURNAL Patent: US 6586572-A 428 01-JUL-2003;
Corixa Corporation; Seattle, WA
FEATURES
  source
    location/Qualifiers
    1. .535
    /organism="unknown"
    /mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCTCTATGGAGGATAC 2442
Db 306 GGCTGGTCTCTATGGAGGATAC 326
RESULT 139
AR453965 535 bp DNA linear PAT 20-FEB-2004
LOCUS
DEFINITION Sequence 428 from patent US 6680197.
ACCESSION AR453965
VERSION AR453965.1 GI:42686755
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 535)
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L.,
Hepler, W.T. and Henderson, R.A.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: US 6680197-A 428 20-JAN-2004;
Corixa Corporation; Seattle, WA
FEATURES
  source
    location/Qualifiers
    1. .535
    /organism="unknown"
    /mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCTCTATGGAGGATAC 2442
Db 306 GGCTGGTCTCTATGGAGGATAC 326
RESULT 140
AR561553 535 bp DNA linear PAT 08-OCT-2004
LOCUS
DEFINITION Sequence 428 from patent US 6756477.
ACCESSION AR561553
VERSION AR561553.1 GI:53974461
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 535)
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J. and Harlocker, S.L.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: US 6756477-A 428 29-JUN-2004;
Corixa Corporation; Seattle, WA
FEATURES
  source
    location/Qualifiers
    1. .535
    /organism="unknown"
    /mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCTCTATGGAGGATAC 2442
Db 306 GGCTGGTCTCTATGGAGGATAC 326
RESULT 137
AR344184 535 bp DNA linear PAT 17-AUG-2003
LOCUS
DEFINITION Sequence 428 from patent US 6579973.
ACCESSION AR344184
VERSION AR344184.1 GI:33740084
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 535)
AUTHORS Yuqiu, J., Dillon, D.C., Mitcham, J.L., Xu, J. and Harlocker, S.L.
TITLE Compositions for the treatment and diagnosis of breast cancer and
methods for their use
JOURNAL Patent: US 6579973-A 428 17-JUN-2003;
Corixa Corporation; Seattle, WA
FEATURES
  source
    location/Qualifiers
    1. .535
    /organism="unknown"
    /mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCTCTATGGAGGATAC 2442
Db 306 GGCTGGTCTCTATGGAGGATAC 326
RESULT 138
AR351385 535 bp DNA linear PAT 17-AUG-2003
LOCUS
DEFINITION Sequence 428 from patent US 6586572.
ACCESSION AR351385
VERSION AR351385.1 GI:33753064
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 535)
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L. and
Hepler, W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
```

```

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 141
AR632008
LOCUS AR632008
DEFINITION Sequence 428 from patent US 6844325.
ACCESSION AR632008
VERSION AR632008.1 GI:59772953
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 535)
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J. and Harlocker, S.L.
TITLE Compositions and methods for the treatment and diagnosis of breast cancer and methods for their use
JOURNAL Patent: US 6844325-A 428 18-JAN-2005;
Corixa Corporation; Seattle, WA
FEATURES
source
1..535
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 142
AR760254
LOCUS AR760254
DEFINITION Sequence 428 from patent US 6958361.
ACCESSION AR760254
VERSION AR760254.1 GI:83327054
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 535)
AUTHORS Houghton, R.L., Sleath, P.R. and Persing, D.H.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: US 6958361-A 428 25-OCT-2005;
Corixa Corporation; Seattle, WA
FEATURES
source
1..535
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 143
AR773397
LOCUS AR773397
DEFINITION Sequence 428 from patent US 6969518.
ACCESSION AR773397
VERSION AR773397.1 GI:83349927
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 535)
AUTHORS Houghton, R.L., Sleath, P.R. and Persing, D.H.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: US 6969518-A 428 29-NOV-2005;
Corixa Corporation; Seattle, WA
FEATURES
source
1..535
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 144
AX303108
LOCUS AX303108
DEFINITION Sequence 428 from Patent WO0179286.
ACCESSION AX303108
VERSION AX303108.1 GI:17383609
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L. and Hepler, W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: WO 0179286-A 428 25-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source
1..535
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 145
BV343295/c
LOCUS BV343295/c
DEFINITION S230F648FC8.T0 Rottweiler Canis familiaris STS genomic, sequence tagged site.
ACCESSION BV343295
VERSION BV343295.1 GI:57543598
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
AUTHORS Canis familiaris (dog)
TITLE Canis familiaris
JOURNAL Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

```

Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.

1 (bases 1 to 588)
Lindblad-Toh, K.
The genome sequence of Canis familiaris
Unpublished (2004)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research

320 Charles Street, Cambridge, MA 02141, USA

Tel: 6172521477

Fax: 6172580903

Email: kersti@genome.wi.mit.edu

Primer A: No sequence submitted

Primer B: No sequence submitted

STS size: 588

Protocol:

WGS-discovery (WGS):
Paired-end low-coverage whole genome shotgun reads were generated
from 9 breeds
(German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador
Retriever, English
Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese
Water Dog -100,000 each)
and five other canids (Chinese, Alaskan, Indian and Spanish Gray
Wolf as well as the
Californian Coyote).

The WGS reads were placed uniquely on the CanFam1.0 boxer assembly
and SNP detection was
carried out by SSAHA-SNP. 863872 reads were annotated as STSs and
485941 SNPs were
annotated with alleles from the boxer and the breed or canid from
which the particular
read came. The validation rate for these SNPs was estimated at
approximately 98%.

A second set of SNPs was generated using a similar methodology
except that the contigs
from the 1.5x poodle assembly (Kirkness 2003) were used instead of
WGS reads. Since this
sequence lacked base quality scores, arbitrary quality scores of
phred 40 were assigned
before the poodle sequence was placed uniquely on the CanFam1.0
boxer assembly and SNP
detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated
with alleles from the
boxer and the poodle. The validation rate for these SNPs was
estimated at approximately 98%.

Internal-WGA-discovery (I-WGA):
A third set of SNPs were discovered by comparing reads in the WGA
assembly. SNPs were
defined as mismatch positions that had a base quality of >= 30 on
both reads in a region
that aligned without gaps, and with at most one additional mismatch
in the ten flanking
bases. For each allele, at least one additional read had to confirm
it. 731476 SNPs were
annotated with alleles between the two boxer alleles. The
validation rate for these SNPs
was estimated at approximately 98%.

Location/Qualifiers

1..588
/organism="Canis familiaris"
/mol_type="Genomic DNA"
/strain="Rottweiler"
/db_xref="taxon:9615"
/map="1 22-531 17237500-17236987"
/clone.lib="Rottweiler"
<1..>588

FEATURES
source

STS
ORIGIN

Query Match 0.7%; Score 21; DB 7; Length 588;

Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3100 AAGATTACTTAAAAA 3120

|||||

Db 71 AAGATTACTTAAAAA 51

RESULT 146

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

1

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..802

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match

Best Local Similarity

Matches

Score 21; DB 2; Length 802;

0.7%; Pred. No. 20;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 GGCTGGTCTTATGGAGGATAC 2442

|||||

Db 326 GGCTGGTCTTATGGAGGATAC 306

RESULT 147

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

1

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..855

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match

Best Local Similarity

Matches

Score 21; DB 2; Length 855;

0.7%; Pred. No. 20;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 GGCTGGTCTTATGGAGGATAC 2442

|||||

```

Db      346  GCCTGGTCTCTATGGAGGATAC 326

RESULT 148
AF007822
LOCUS   Homo sapiens cytoplasmic Sepsis truncated isoform mRNA, complete cds.
DEFINITION
ACCESSION AF007822      1143 bp      mRNA      linear      PRI 27-JAN-2000
VERSION   AF007822.1  GI:6649854
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE
AUTHORS   Goldstein,L.A. and Chen,W.-T.
TITLE     Identification of an alternatively spliced sepsis mRNA that encodes a novel intracellular isoform
JOURNAL   J. Biol. Chem. 275 (4), 2554-2559 (2000)
PUBMED    10644713
AUTHORS   2 (bases 1 to 1143)
TITLE     Goldstein,L.A. and Chen,W.-T.
JOURNAL   Direct Submission
SUBMITTED (11-JUN-1997) Lombardi Cancer Center, Georgetown University Medical Center, 3970 Reservoir Road NW, Washington, DC 20007-2187, USA
FEATURES
source    Location/Qualifiers
1..1143
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_lines="LOX"
/cell_type="melanoma"
389..1108
/note="fibroblast activation protein; intracellular isoform equivalent to carboxyl-terminal region of serine integral membrane protease Sepsis; alternative splice variant"
CDS
/codon_start=1
/product="cytoplasmic Sepsis truncated isoform"
/protein_id="QAF3160.1"
/db_xref="GI:664985"
/translation="MILFFQDRSKKYPILLIQVGGPCSVSRVPAVNWISYLASKE
GMVIALVDGRTAQGDKLIYAVTRKLGVEVEDQITAVRKFIEMGFIDEKRIATWGW
SYGCVASLALASGTLGKCGIAVAPVSSWEYVYTFERFGLPTKDDNLEHYKNST
VMARAEVRNVDILLIHGTADDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGLSL
STNHLVTHMTHFLKQCFSLSD"

ORIGIN
Query Match      0.7%; Score 21; DB 5; Length 1143;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 GCCTGGTCTCTATGGAGGATAC 2442
|||||
Db      689  GCCTGGTCTCTATGGAGGATAC 709

RESULT 149
HSM803473
LOCUS   Homo sapiens mRNA; cDNA DKFZp686G13158 (from clone DKFZp686G13158).
DEFINITION
ACCESSION AL832166
VERSION   AL832166.1  GI:21732710
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE
AUTHORS   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1615)

AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
TITLE     Direct Submission
JOURNAL   Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherrberg, GERMANY
COMMENT    Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp686G13158) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
source    Location/Qualifiers
1..1615
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cldones="DKFZp686G13158"
/tissue_type="human colon endothel primary cell culture"
/cldone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
1513..1518
polyA_site
polyA_site 1536
ORIGIN
Query Match      0.7%; Score 21; DB 5; Length 1615;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 GCCTGGTCTCTATGGAGGATAC 2442
|||||
Db      823  GCCTGGTCTCTATGGAGGATAC 843

RESULT 150
CS141101
LOCUS   Sequence 3 from Patent WO2005071073.
DEFINITION CS141101
ACCESSION  CS141101
VERSION    CS141101.1  GI:73531315
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS   McLean,P.A.
TITLE     Pap compositions and the use thereof for immunomodulation
JOURNAL   Patent WO 2005071073-A 3 04-AUG-2005;
Point Therapeutics, Inc. (US)
FEATURES
source    Location/Qualifiers
1..2255
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      0.7%; Score 21; DB 2; Length 2255;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 GCCTGGTCTCTATGGAGGATAC 2442
|||||
Db      1771 GCCTGGTCTCTATGGAGGATAC 1791

Search completed: June 22, 2006, 08:46:52
Job time : 17308 secs

```

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:



BLACK BORDERS



IMAGE CUT OFF AT TOP, BOTTOM OR SIDES



FADED TEXT OR DRAWING



BLURRED OR ILLEGIBLE TEXT OR DRAWING



SKEWED/SLANTED IMAGES



COLOR OR BLACK AND WHITE PHOTOGRAPHS



GRAY SCALE DOCUMENTS



LINES OR MARKS ON ORIGINAL DOCUMENT



REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY



OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)